EDITORIAL BOARD

Editor

Barney Warf
University of Kansas

Associate Editors

Piotr Jankowski
San Diego State University

Barry D. Solomon
Michigan Technological University

Mark Welford
Georgia Southern University

Managing Editor

Jonathan Leib
Old Dominion University
## CONTENTS

<table>
<thead>
<tr>
<th>Entries</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>536</td>
</tr>
<tr>
<td>D</td>
<td>663</td>
</tr>
<tr>
<td>E</td>
<td>809</td>
</tr>
</tbody>
</table>

List of Entries

vii
LIST OF ENTRIES

Abler, Ronald
Absolute Space
Accessibility
Acid Rain
Actor-Network Theory
Adaptation to Climate Change
Adaptive Harvest Management
Adaptive Radiation
Adiabatic Temperature Changes
Aerial Imagery: Data
Aerial Imagery: Interpretation
African Union
Agamben, Giorgio
Agent-Based Models
Agglomeration Economies
Aging and the Aged, Geography of. See Elderly, Geography and the
Agnew, John
Agricultural Biotechnology
Agricultural Intensification
Agricultural Land Use
Agriculture, Industrialized
Agriculture, Preindustrial
Agrobiodiversity
Agrochemical Pollution
Agroecology
Agrofoods
Agroforestry
AIDS, Geography of. See Disease, Geography of; HIV/AIDS, Geography of
Air Masses
Air Pollution. See Atmospheric Pollution
Albedo
al-Idrisi
Altitude
Ambient Air Quality
American Geographical Society
Analytical Operations in GIS
Anarchism and Geography
Anaximander
Animal Geographies
Annales School
Anselin, Luc
Antevs, Ernst
Anthropogenic Atmospheric Change. See Anthropogenic Climate Change;
Stratospheric Ozone Depletion
Anthropogenic Climate Change
Anthropogeography
Antiglobalization
Antipodes
Antisystemic Movements
Applied Geography
Appropriate Technology. See Sustainable Development; Sustainable Development Alternatives; Sustainable Production
Aquaculture
Archipelago
Architecture and Geography Argumentation Maps
Arid Topography
Aristotle
Armstrong, Marc
Art and Geography
Asia-Pacific Economic Cooperation
Association of American Geographers
Association of Geographic Information Laboratories for Europe
Association of Southeast Asian Nations
Atmospheric Circulation
Atmospheric Composition and Structure
Atmospheric Energy Transfer
Atmospheric Moisture
Atmospheric Particulates Across Scales
Atmospheric Pollution
Atmospheric Pressure
Atmospheric Remote Sensing
Atmospheric Variations in Energy
Atoll
Automobile Industry
Automobility
Avalanches
Aviation and Geography
Barrier Islands
Barrows, Harlan
Basin and Range Topography
<table>
<thead>
<tr>
<th>Entry</th>
<th>Entry</th>
<th>Entry</th>
</tr>
</thead>
<tbody>
<tr>
<td>Batty, Michael</td>
<td>Bayesian Statistics in Spatial Analysis</td>
<td>Class, Geography and Class, Nature and</td>
</tr>
<tr>
<td>Behavioral Geography</td>
<td>Berkeley School</td>
<td>Client-Server Architecture</td>
</tr>
<tr>
<td>Berry, Brian</td>
<td>Bhopal, India, Chemical Disaster</td>
<td>Climate: Dry</td>
</tr>
<tr>
<td>Bhopal, India, Chemical Disaster</td>
<td>Biblical Mapping</td>
<td>Climate: Midlatitude, Mild</td>
</tr>
<tr>
<td>Biographical Geography</td>
<td>Biofuels</td>
<td>Climate: Midlatitude, Severe</td>
</tr>
<tr>
<td>Biogeography</td>
<td>Biogeochemical Cycles</td>
<td>Climate: Mountain</td>
</tr>
<tr>
<td>Biome: Boreal Forest</td>
<td>Biome: Desert</td>
<td>Climate: Polar</td>
</tr>
<tr>
<td>Biome: Desert</td>
<td>Biome: Midlatitude Deciduous Forest</td>
<td>Climate: Tropical Humid</td>
</tr>
<tr>
<td>Biome: Midlatitude Grassland</td>
<td>Biome: Tropical Deciduous Forest</td>
<td>Climate Change</td>
</tr>
<tr>
<td>Biome: Tropical Rain Forest</td>
<td>Biome: Tropical Savanna</td>
<td>Climate Policy</td>
</tr>
<tr>
<td>Biome: Tropical Scrub</td>
<td>Biome: Tundra</td>
<td>Climate Types</td>
</tr>
<tr>
<td>Biome: Tundra</td>
<td>Biophysical Remote Sensing</td>
<td>Climates Rict</td>
</tr>
<tr>
<td>Bioregionalism</td>
<td>Biosphere Reserves</td>
<td>Climatology</td>
</tr>
<tr>
<td>Biota and Climate</td>
<td>Biota and Soils</td>
<td>Clouds</td>
</tr>
<tr>
<td>Biota and Soils</td>
<td>Biota and Topography</td>
<td>Clusters</td>
</tr>
<tr>
<td>Biota Migration and Dispersal</td>
<td>Biotechnology and Ecological Risk</td>
<td>Coal</td>
</tr>
<tr>
<td>Biotechnology Industry</td>
<td>Chorological Risk</td>
<td>Coastal Dead Zones</td>
</tr>
<tr>
<td>Biruni</td>
<td>Choropleth Maps</td>
<td>Coastal Erosion and Deposition</td>
</tr>
<tr>
<td>Blaikie, Piers</td>
<td>Chinese Spills, Environment and Society</td>
<td>Coastal Hazards</td>
</tr>
<tr>
<td>Blaut, James</td>
<td>Chlorofluorocarbons</td>
<td>Coastal Zone and Marine Pollution</td>
</tr>
<tr>
<td>Blindness and Geography</td>
<td>Chimneys/Foehns</td>
<td>Cold War, Geography of</td>
</tr>
<tr>
<td>Body, Geography of Borderlands</td>
<td>Chow, Richard</td>
<td>Collaborative GIS</td>
</tr>
<tr>
<td>Borders and Boundaries</td>
<td>Chorology</td>
<td>Colonialism</td>
</tr>
<tr>
<td>Bourghassi, Carmina</td>
<td>Choropleth Maps</td>
<td>Color in Map Design</td>
</tr>
<tr>
<td>Bowman, Isaiah</td>
<td>Chinese Spills, Environment and Society</td>
<td>Columbus, Christopher</td>
</tr>
<tr>
<td>Brownfields</td>
<td>Chlorinated Hydrocarbons</td>
<td>Commodity Chains</td>
</tr>
<tr>
<td>Built Environment</td>
<td>Chlorofluorocarbons</td>
<td>Common Pool Resources</td>
</tr>
<tr>
<td>Bunge, William</td>
<td>Chorolera, Geography of</td>
<td>Common Property Resource Management</td>
</tr>
<tr>
<td>Business Cycles and Geography</td>
<td>Chorley, Richard</td>
<td>Commons, Tragedy of the</td>
</tr>
<tr>
<td>Business Geography</td>
<td>Chinooks/Foehns</td>
<td>Commonwealth of Independent States</td>
</tr>
<tr>
<td>Business Models for Geographic Information Systems</td>
<td>Christine, Nicholas</td>
<td>Communications Geography</td>
</tr>
<tr>
<td>Buttimmer, Anne</td>
<td>Circassar, Walter</td>
<td>Communism and Geography</td>
</tr>
<tr>
<td>Cadastral Systems</td>
<td>Circuits of Capital</td>
<td>Community-Based Conservation</td>
</tr>
<tr>
<td>CAD Systems</td>
<td>Citizenship</td>
<td>Community-Based Environmental Planning</td>
</tr>
<tr>
<td>Câmara, Gilberto</td>
<td>Clark, Andrew</td>
<td>Community-Based Resource Management</td>
</tr>
<tr>
<td>Canadian Association of Geographers</td>
<td>Clark, William</td>
<td>Environmental Planning</td>
</tr>
<tr>
<td>Cancer, Geography of Carbonation</td>
<td>Common Property Resource Management</td>
<td>Community-Based Resource Management</td>
</tr>
<tr>
<td>Carbon Cycle</td>
<td>Commons, Tragedy of the</td>
<td>Environmental Planning</td>
</tr>
<tr>
<td>Carbon Trading and Carbon Offsets</td>
<td>Commonwealth of Independent States</td>
<td>Community Forestry</td>
</tr>
<tr>
<td>Carcinogens</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Commuting</td>
</tr>
<tr>
<td>Carrying Capacity</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Comparative Advantage</td>
</tr>
<tr>
<td>Cartograms</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Competitive Advantage</td>
</tr>
<tr>
<td>Cartography</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Complexity Theory</td>
</tr>
<tr>
<td>Cartography, History of Castells, Manuel</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Complex Systems Models</td>
</tr>
<tr>
<td>Caverns</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Cellular Automata</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Census</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Census Tracts</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Centers of Domestication</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Central Business District</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Central Place Theory</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Chemical Spills, Environment, and Society</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Chernobyl Nuclear Accident</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Chicago School</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Childhood Spatial and Environmental Learning</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Children, Geography of Chinook/Foehns</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Chipko Movement</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Chlorinated Hydrocarbons</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Chlorofluorocarbons</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Cholera, Geography of</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Chorley, Richard</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Chorology</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Choropleth Maps</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Business Cycles and Geography</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Business Geography</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Class, Geography and Class, Nature and Client-Server Architecture</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Climate: Dry</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Climate: Midlatitude, Mild</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Climate: Midlatitude, Severe</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Climate: Mountain</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Climate: Polar</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Climate: Tropical Humid</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Climate Change</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Climate Policy</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Climate Types</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Climatic Rict</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Climatology</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Clouds</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Clusters</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Coal</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Coastal Dead Zones</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Coastal Erosion and Deposition</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Coastal Hazards</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Coastal Zone and Marine Pollution</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Cold War, Geography of</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Collaborative GIS</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Colonialism</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Color in Map Design</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Columbus, Christopher</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Commodity Chains</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Common Pool Resources</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Common Property Resource Management</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Commons, Tragedy of the</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Commonwealth of Independent States</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Communications Geography</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Communication and Geography</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Community-Based Conservation</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Community-Based Environmental Planning</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Community-Based Resource Management</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Environmental Planning</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Community-Based Natural Resource Management</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Resource Management</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Community Forestry</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Commuting</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Comparative Advantage</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Competitive Advantage</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Complexity Theory</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Complex Systems Models</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
<tr>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
<td>Conference of Latin Americanist Geographers</td>
</tr>
</tbody>
</table>
Conflation
Conservation
Conservation Zoning
Consumption, Geographies of
Continental Drift. See Plate Tectonics
Cook, Captain James
Coordinate Geometry
Coordinate Systems
Coordinate Transformations
Coral Reef
Coral Reef Geomorphology
Coriolis Force
Corporate Voluntary Environmental Initiatives and Self-Regulation
Cosgrove, Denis
Cosmopolitanism
Cost-Benefit Analysis
Council for Mutual Economic Assistance (COMECON)
Countermapping
Counterurbanization
Coupled Human and Animal Systems
Coupled Human and Natural Systems
Creep
Crime, Geography of Crisis
Critical Geopolitics
Critical GIS
Critical Human Geography
Critical Studies of Nature
Crop Genetic Diversity
Crop Rotation
Cross-Border Cooperation
Cultural Ecology
Cultural Geography
Cultural Landscape
Cultural Turn
Cyberspace
Cyborg Ecologies
Cyclones: Extratropical Cyclones: Occluded
Dangermond, Jack
Darby, Henry Clifford
Darwinism and Geography
Dasymetric Maps
Database Management Systems
Database Versioning
Data Classification Schemes
Data Compression Methods
Data Editing
Data Format Conversion
Data Indexing
Data Querying in GIS
Datums
Davis, William Morris
Dear, Michael
Debt and Debt Crisis
Decolonization
Deep Ecology Movements
Deforestation
Deindustrialization
Delta
Democracy
Demographic Transition
Dendrochronology
Dependency Theory
Derechos
Desertification
Desert Varnish
Deterritorialization and Reterritorialization
Developing World
Development Theory
Diamond, Jared
Diaspora
Diastrophism
Difference, Geographies of Differential Heating
Differential Vulnerabilities to Hazards
Diffusion
Digital Divide
Digital Terrain Model
Digitizing
Disability, Geography of
Disaster Prediction and Warning
Disaster Preparedness
Discourse and Geography
Disease, Geography of
Distance Decay
Distributed Computing
Distribution of Resource Access
Division of Labor
Domestication Centers. See Centers of Domestication
Domestication of Animals
Domestication of Plants
Domino Theory
Dot Density Maps
Drought Risk and Hazard
Drugs, Geography of Dunes
Dynamic and Interactive Displays
Earle, Carville
Earthquakes
Earth’s Coordinate Grid
Eastman, Ronald
Ecofeminism
Ecological Economics
Ecological Fallacy
Ecological Footprint
Ecological Imaginaries
Ecological Justice
Ecological Mapping
Ecological Modernization
Ecological Regimes
Ecological Risk Analysis
Ecological Services. See Environmental Services
Ecological Zones
E-Commerce and Geography
Economic Base Analysis
Economic Geography
Economies of Scale
Economies of Scope
Ecoregions
Ecoshed
Ecosystem Decay
Ecosystems
Ecotone
Ecotourism
Education, Geographies of
Egenhofer, Max
Elderly, Geography and the
Electoral Geography
Electronic Atlases
Electronics Industry, Geography of
El Niño
Emerging Markets
| Emotions, Geography and Empiricism | Environmental Protection |
| Endogenous Growth Theory. See Knowledge Spillovers; Learning Regions | Environmental Racism |
| Energy and Human Ecology | Environmental Refugees |
| Energy Models | Environmental Restoration |
| Energy Policy | Environmental Rights |
| Energy Resources | Environmental Security |
| Enlightenment | Environmental Services |
| Enterprise GIS | Environmental Social Movements. See International Environmental Movements |
| Environmental Certification | Environment and Development |
| Environmental Determinism | Epistemology |
| Environmental Discourse | Equator |
| Environmental Entitlements | Equinox |
| Environmental Ethics | Eratosthenes |
| Environmental Footprint. See Ecological Footprint | Error Propagation |
| Environmental History | Ethics, Geography and Ethnicity |
| Environmental Imaginaries | Ethnicity |
| Environmental Impact Assessment | Ethnicity and Nature |
| Environmental Impacts of Agriculture | Ethnic Segregation |
| Environmental Impacts of Cities | Ethnocentrism |
| Environmental Impacts of Manufacturing | Eurocentrism |
| Environmental Impacts of Mining. See Open-Pit Mining | Euromarket |
| Environmental Impacts of Oil Fields | European Green Movements |
| Environmental Impacts of Pipelines | European Union |
| Environmental Impacts of Roads | Everglades Restoration |
| Environmental Impacts of Tourism | Everyday Life, Geography and Existentialism and Geography |
| Environmental Impacts of War | Exotic Species |
| Environmental Impact Statement | Exploration |
| Environmental Justice | Exploratory Spatial Data Analysis |
| Environmental Law | Export-Led Development |
| Environmental Management | Export Processing Zones |
| Environmental Management: Drylands | Externalities |
| Environmental Mapping | Extinctions |
| Environmental Perception | Extractive Reserves |
| Environmental Planning | Extinctions |
| Feminist Environmentalism | Extreme Geography |
| Feminist Geographies | Exurbs |
| Feminist Methodologies | Factors Affecting Location of Firms |
| Feminist Political Ecology | Fair Trade and Environmental Certification |
| Fertility Rate | Famine, Geography of Faulting |
| Fieldwork in Human Geography | Fear, Geographies of Fevre, Lucien |
| Fieldwork in Physical Geography | Feminist Environmental Geographies |
| Film and Geography | Filtering |
| Filtration | Finance, Geography of Fisher, Peter |
| Finance, Geography of Fisher, Peter | Fish Farming |
| Fjords | Flash Floods |
| Floodplain | Floods |
| Floods | Flow |
| Flow | Flow Maps |
| Flow Maps | Folding |
| Folding | Folk Culture and Geography |
| Folk Culture and Geography | Food, Geography of Food and Agriculture Organization (FAO) |
| Footprint Analysis. See Ecological Footprint | Footprint Analysis. See Ecological Footprint |
| Fordism | Foreign Aid |
| Foreign Aid | Foreign Direct Investment |
| Foreign Direct Investment | Forest Degradation |
| Forest Degradation | Forest Fragmentation |
| Forest Fragmentation | Forest Land Use |
| Forest Land Use | Forest Restoration |
| Forest Restoration | Fotheringham, A. Stewart |
| Fotheringham, A. Stewart | Frank, Andrew |
| Frank, Andrew | Frontiers |
| Frontiers | Fronts |
| Fronts | Gaia Theory |
| Gaia Theory | Gama, Vasco da |
| Gama, Vasco da | Game Ranching |
| Game Ranching | Gated Community |
| Gated Community | Gays and Lesbians, Geography and/of Gazetteers |
LIST OF ENTRIES

Gender and Environmental Hazards
Gender and Geography
Gender and Nature
General Circulation Model (GCM). See Anthropogenic Climate Change; Atmospheric Energy Transfer; Climate Change
Genetically Modified Organisms (GMOs)
Genocide, Geographies of
Gentrification
Geocoding
Geocomputation
Geodemographics
Geodesy
Geographical Ignorance
Geographical Imagination
Geographically Weighted Regression
Geographic Information Systems
Geography Education
Geolibraries
Geologic Timescale
Geomancy
Geometric Correction
Geometric Measures
Geomorphologic Cycle
Geomorphology
Geophagy
Geopolitics
Geosensor Networks
Geospatial Industry
Geospatial Semantic Web
Geostatistics
Geothermal Energy
Geothermal Features
Geovisualization. See Cartography; Dynamic and Interactive Displays; Three-Dimensional Models
Getis, Arthur
Ghetto
Giddens, Anthony
Gilbert, Grove Karl
GIS, Environmental Model Integration and
GIS, History of GIScience
GIS Design
GIS Implementation
GIS in Archaeology
GIS in Disaster Response
GIS in Environmental Management
GIS in Health Research and Health Care
GIS in Land Use Management
GIS in Local Government
GIS in Public Policy
GIS in Transportation
GIS in Urban Planning
GIS in Utilities
GIS in Water Management
GIS Software
GIS Web Services
Glaciers: Continental
Glaciers: Mountain
Global Climate Change. See Anthropogenic Climate Change
Global Environmental Change
Globalization
Global Positioning System
Global Sea-Level Rise
Global Warming. See Anthropogenic Climate Change
Globes. See Cartography; Map Projections
Glocalization
Golledge, Reginald
Goodchild, Michael
Goode, J. Paul
Google Earth
Gottmann, Jean
Gould, Peter
Governance
Governmentality and Conservation
Gravity Model
Great American Exchange
Greenbelts
Green Building
Green Design and Development
Greenhouse Gases
Green Revolution. See Environmental Impacts of Agriculture
Gregory, Derek
Gross Domestic Product/Gross National Product
Ground Reference Data
Groundwater
Growth Machine
Growth Poles
Gully Erosion
Guyot, Arnold
Hadley Cell
Hägerstrand, Torsten
Haggett, Peter
Hanson, Susan
Harley, Brian
Hartshorne, Richard
Harvey, David
Hate, Geographies of
Haushofer, Karl
Hayden, Ferdinand
Health and Health Care, Geography of
Heavy Metals as Pollutants
Hegemony
Herbicides
Herodotus
Hettner, Alfred
Heuristic Methods in Spatial Analysis
High-Performance Computing
High Technology
Hipparchus
Historical Geography
Historicism
Historic Preservation
History of Geography. See Cartography, History of; GIS, History of; Human Geography, History of; Physical Geography, History of
HIV/AIDS, Geography of
Home
Homelessness
Hou Renzhi
Housing and Housing Markets
Housing Policy
Hoyt, Homer
Human Dimensions of Global Environmental Change
Human Ecology
Human Geography, History of
Human-Induced Invasion of Species
Humanistic Geography
Humanistic GIScience
Human Rights, Geography and Humboldt, Alexander von
Humidity
Hunger
Hunting and Gathering
Huntington, Ellsworth
Hurricane Katrina
Hurricanes, Physical Geography of
Hurricanes, Risk and Hazard
Hybrid Geographies
Hybridization of Plant and Animal Species
Hydroelectric Power
Hydrological Connectivity
Hydrology
Hydrothermal Energy. See Geothermal Energy
Ibn Battuta
Ibn Khaldūn
Ice
Identity, Geography and
Idiographic
Image Enhancement
Image Fusion
Image Interpretation
Image Processing
Image Registration
Image Texture
Imaging Spectroscopy
Immigration
Imperialism
Impermeable Surfaces
Import Substitution
Industrialization
Incubator Zones
Indigeneity
Indigenous Agriculture
Indigenous and Community Conserved Areas
Indigenous Cartographies
Indigenous Environmental Knowledge
Indigenous Environmental Practices
Indigenous Forestry
Indigenous Reserves
Indigenous Water Management
Industrial Districts
Industrial Ecology
Industrialization
Industrial Revolution
Inequality and Geography
Informal Economy
Information Society
Infrastructure
Innovation, Geography of
Input-Output Models
Institute of British Geographers
Interactive Mapping. See Map Animation
Intergovernmental Environmental Organizations and Initiatives
International Criminal Court
International Environmental Movements
International Environmental NGOs
International Geographical Union
International Monetary Fund
International Watershed Management
Internet. See Communications, Geography of; Cyberspace; Digital Divide; Telecommunications and Geography
Internet GIS
Interoperability and Spatial Data Standards
Interviewing
Invasion and Succession
Isard, Walter
Island Biogeography
Islands, Small
Isopleth Maps
Jackson, John Brinckerhoff
Jefferson, Thomas
Johnston, R. J.
Journey-to-Work. See Commuting
Justice, Geography of
Kant, Immanuel
Karst Topography
Kates, Robert
Keystone Species
Knowledge, Geography of
Knowledge Spillovers
Köppen, Wladimir
Köppen-Geiger Climate Classification
Kropotkin, Peter
Krummholtz
Kuhn, Werner
Kwan, Mei-Po
Kyoto Protocol. See Climate Policy
Labor, Geography of
Lahar. See Volcanoes
Land Degradation
Landfills
Landforms
Land Reform
Land Use Analysis
Land Use and Cover Change (LUCC)
Land Use and Land Cover Mapping
Land Use History
Land Use Planning
<table>
<thead>
<tr>
<th>Land-Water Breeze</th>
<th>Marcus, Melvin G.</th>
<th>Multivariate Mapping</th>
</tr>
</thead>
<tbody>
<tr>
<td>Languages, Geography of La Niña</td>
<td>Marginal Regions</td>
<td>Music and Sound, Geography and</td>
</tr>
<tr>
<td>Lapse Rate</td>
<td>Marine Aquaculture</td>
<td>Nation</td>
</tr>
<tr>
<td>Latent Heat</td>
<td>Maritime Spaces. See Oceans</td>
<td>National Aeronautics and Space Administration (NASA)</td>
</tr>
<tr>
<td>Latitude</td>
<td>Mark, David M.</td>
<td>National Center for Geographic Information and Analysis</td>
</tr>
<tr>
<td>Law, Geography of Learning Regions</td>
<td>Market-Based Environmental Regulation</td>
<td>National Council for Geographic Education</td>
</tr>
<tr>
<td>Lefebvre, Henri</td>
<td>Marsh, George Perkins</td>
<td>National Geographic Society</td>
</tr>
<tr>
<td>Legal Aspects of Geospatial Information</td>
<td>Marxism, Geography and Masculinities and Geography</td>
<td>Nationalism</td>
</tr>
<tr>
<td>Lewis, Peirce</td>
<td>Massey, Doreen</td>
<td>Natural Growth Rate</td>
</tr>
<tr>
<td>Lewis and Clark Expedition</td>
<td>Mass Wasting</td>
<td>Natural Hazards and Risk Analysis</td>
</tr>
<tr>
<td>Ley, David</td>
<td>Mather, John Russell</td>
<td>Nature</td>
</tr>
<tr>
<td>LiDAR and Airborne Laser Scanning</td>
<td>Maury, Matthew Fontaine</td>
<td>Nature-Society Theory</td>
</tr>
<tr>
<td>Lightning</td>
<td>McKnight, Tom L.</td>
<td>Neighborhood</td>
</tr>
<tr>
<td>Lillesand, Thomas</td>
<td>Media and Geography</td>
<td>Neocolonialism</td>
</tr>
<tr>
<td>Linear Referencing and Dynamic Segmentation</td>
<td>Medical Geography</td>
<td>Neogeography</td>
</tr>
<tr>
<td>Literature, Geography and Livingstone, David</td>
<td>Meinig, Donald</td>
<td>Neoliberal Environmental Policy</td>
</tr>
<tr>
<td>Locality</td>
<td>Mental Maps</td>
<td>Neoliberalism</td>
</tr>
<tr>
<td>Locally Unwanted Land Uses (LULUs)</td>
<td>Mercator, Gerardus</td>
<td>Neo-Malthusianism</td>
</tr>
<tr>
<td>Location-Allocation Modeling</td>
<td>Metadata</td>
<td>Network Analysis</td>
</tr>
<tr>
<td>Location-Based Services</td>
<td>Metropolitan Area</td>
<td>Network Data Model</td>
</tr>
<tr>
<td>Location Quotients</td>
<td>Microwave/RADAR Data</td>
<td>New International Division of Labor</td>
</tr>
<tr>
<td>Location Theory</td>
<td>Migration</td>
<td>Newly Industrializing Countries</td>
</tr>
<tr>
<td>Logical Positivism</td>
<td>Military Geography</td>
<td>New Urbanism</td>
</tr>
<tr>
<td>Longitude</td>
<td>Military Spending</td>
<td>Nitrogen Cycle</td>
</tr>
<tr>
<td>Los Angeles School</td>
<td>Miller, Harvey J.</td>
<td>Nomadic Herding</td>
</tr>
<tr>
<td>Lösch, August</td>
<td>Minerals</td>
<td>Nomadism</td>
</tr>
<tr>
<td>Love Canal</td>
<td>Mining and Geography</td>
<td>Nomothetic</td>
</tr>
<tr>
<td>Lynch, William</td>
<td>Mitchell, Don</td>
<td>Nongovernmental Organizations (NGOs)</td>
</tr>
<tr>
<td>MacEachren, Alan</td>
<td>Mixed Farming</td>
<td>Nonpoint Sources of Pollution</td>
</tr>
<tr>
<td>Mackinder, Sir Halford</td>
<td>Mobile GIS</td>
<td>Nonrenewable Resources</td>
</tr>
<tr>
<td>Magellan, Ferdinand</td>
<td>Mobility</td>
<td>Nonrepresentational Theory</td>
</tr>
<tr>
<td>Mahan, Alfred Thayer</td>
<td>Models and Modeling</td>
<td>Nonvisual Geographies</td>
</tr>
<tr>
<td>Malaria, Geography of Malthusianism</td>
<td>Modernity</td>
<td>North American Free Trade Agreement (NAFTA)</td>
</tr>
<tr>
<td>Manufacturing Belt</td>
<td>Modernization Theory</td>
<td>North Atlantic Treaty Organization (NATO)</td>
</tr>
<tr>
<td>Map Algebra</td>
<td>Modifiable Areal Unit Problem</td>
<td>Not in My Backyard (NIMBY)</td>
</tr>
<tr>
<td>Map Animation</td>
<td>Money, Geographies of Monmonier, Mark</td>
<td>Nuclear Energy</td>
</tr>
<tr>
<td>Map Design</td>
<td>Monsoons</td>
<td>Nonrepresentational Theory</td>
</tr>
<tr>
<td>Map Evaluation and Testing</td>
<td>Morrill, Richard</td>
<td>Nonvisual Geographies</td>
</tr>
<tr>
<td>Map Generalization</td>
<td>Morse, Jedediah</td>
<td>North American Free Trade Agreement (NAFTA)</td>
</tr>
<tr>
<td>Map Projections</td>
<td>Mortality Rate</td>
<td>North Atlantic Treaty Organization (NATO)</td>
</tr>
<tr>
<td>Map Visualization</td>
<td>Most Favored Nation Status</td>
<td>Not in My Backyard (NIMBY)</td>
</tr>
<tr>
<td></td>
<td>Movimento Sem Terra</td>
<td>Nuclear Energy</td>
</tr>
</tbody>
</table>
Nutrient Cycles
Nyerges, Timothy

Object-Based Image Analysis
Oceanic Circulation
Oceans
Offshore Finance
Oil Spills
Okabe, Atsuyuki
Olsson, Gunnar
Ontological Foundations of Geographical Data
Ontology
Open Geodata Standards
Open Geospatial Consortium (OGC)
Open-Pit Mining
Open Source Geospatial Foundation
Open Source GIS
Open Space
Organic Agriculture
Organisation for Economic Co-operation and Development (OECD)
Organization of the Petroleum Exporting Countries (OPEC)
Organophosphates
Orientalism
Ortelius
Other/Otherness
Outsourcing
Overpopulation. See Malthusianism; Neo-Malthusianism

Palimpsest
Panchromatic Imagery
Panopticon
Parks and Reserves
Parsons, James
Participant Observation
Participatory Learning and Action
Participatory Mapping
Participatory Planning
Participatory Rural Appraisal
Pastoral Herding. See Nomadic Herding

Patches and Corridors in Wildlife Conservation
Path Dependence
Patriarchy, Geography and Peasants and Peasantry
Peat
Pedology. See Soils
Peet, Richard
Penck, Walther
Periglacial Environments
Permaculture
Permafrost
Pesticides
Pest Management
Petroleum
Peuquet, Donna
Phenomenology
Phosphorus Cycle
Photochemical Smog
Photogrammetric Methods
Photography, Geography and Physical Geography, History of
Pickles, John
Pilgrimage
Place
Place Names
Place Promotion
Plantations
Plate Tectonics
Playas
Point Pattern Analysis
Point Sources of Pollution
Poles, North and South
Political Ecology
Political Economy
Political Economy of Resources
Political Geography
Polychlorinated Biphenyls (PCBs)
Popular Culture, Geography and Population and Land
Degradation
Population and Land Use
Population Density
Population, Environment, and Development
Population Geography
Population Pyramid
Portolan Charts
Ports and Maritime Trade

Positionality
Positivism. See Logical Positivism
Postcolonialism
Post-Fordism. See Flexible Production
Postindustrial Society
Postmodernism
Poststructuralism
Poverty
Powell, John Wesley
Prairie Restoration
Prairies
Precipitation, Global
Precipitation Formation
Pred, Allan
Primate Cities
Prime Meridian. See Longitude
Privacy and Security of Geospatial Information
Producer Services
Product Cycle
Production of Space
Psychoanalysis, Geography and Ptolemy
Public Housing
Public Participation GIS
Public Policy, Geography of Public-Private Partnerships
Public Space
Public Water Services
Pyrogeography

Qualitative Methods
Quantitative Methods
Quantitative Revolution
Queer Theory

Race and Empire
Race and Nature
Race and Racism
Racial Segregation
Radiation: Solar and Terrestrial
Radical Geography
Radiometric Correction
Radiometric Normalization
Radiometric Resolution
Railroads and Geography
Raisz, Erwin
<table>
<thead>
<tr>
<th>Steel Industry, Geography of</th>
<th>Textile Industry</th>
<th>United Nations Conference on Environment and Development</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stereoscopy and Orthoimagery</td>
<td>Thales</td>
<td>United Nations Environmental Summits</td>
</tr>
<tr>
<td>Storper, Michael</td>
<td>Thermal Imagery</td>
<td>United Nations Environment Programme (UNEP)</td>
</tr>
<tr>
<td>Strabo</td>
<td>Thorntwaite, C. Warren</td>
<td>United States Censuses Bureau</td>
</tr>
<tr>
<td>Strahler, Arthur</td>
<td>Three-Dimensional Data Models</td>
<td>United States Geological Survey (USGS)</td>
</tr>
<tr>
<td>Stratospheric Ozone Depletion</td>
<td>Three Mile Island Nuclear Accident</td>
<td>University Consortium for Geographic Information Science</td>
</tr>
<tr>
<td>Strip Mining</td>
<td>Thrift, Nigel</td>
<td>Unsupervised Classification</td>
</tr>
<tr>
<td>Structural Adjustment</td>
<td>Thunderstorms</td>
<td>Unwin, David</td>
</tr>
<tr>
<td>Structuralism</td>
<td>Thuinen Model</td>
<td>Urban and Regional Development</td>
</tr>
<tr>
<td>Structuration Theory</td>
<td>Timber Plantations</td>
<td>Urban and Regional Planning</td>
</tr>
<tr>
<td>Subaltern Studies</td>
<td>Time, Geographies of</td>
<td>Urban Ecology</td>
</tr>
<tr>
<td>Suburban Land Use</td>
<td>Time-Geography</td>
<td>Urban and Regional Planning</td>
</tr>
<tr>
<td>Suburbs and Suburbanization</td>
<td>Time-Space Compression</td>
<td>Urban Sustainability</td>
</tr>
<tr>
<td>Sui, Daniel</td>
<td>T-in-O Maps</td>
<td>Urban Land Use</td>
</tr>
<tr>
<td>Suitability Analysis</td>
<td>Tobler, Waldo</td>
<td>Urban Metabolism</td>
</tr>
<tr>
<td>Sunbelt</td>
<td>Tomlinson, Roger</td>
<td>Urban Planning and Geography</td>
</tr>
<tr>
<td>Supervised Classification</td>
<td>Topological Relationships</td>
<td>Urban Policy</td>
</tr>
<tr>
<td>Supranational Integration</td>
<td>Toponomy</td>
<td>Urban Solid Waste Management</td>
</tr>
<tr>
<td>Surface Water</td>
<td>Topophilia</td>
<td>Urban Sprawl</td>
</tr>
<tr>
<td>Surveillance</td>
<td>Tornadoes</td>
<td>Urban Storm Water Management</td>
</tr>
<tr>
<td>Surveying</td>
<td>Tourism</td>
<td>Urban Sustainability</td>
</tr>
<tr>
<td>Sustainability Science</td>
<td>Township and Range System</td>
<td>Urban Underclass</td>
</tr>
<tr>
<td>Sustainable Agriculture</td>
<td>Trade</td>
<td>Urban Water Supply</td>
</tr>
<tr>
<td>Sustainable Cities</td>
<td>Transhumance. See Nomadism</td>
<td>Usability of Geospatial Information</td>
</tr>
<tr>
<td>Sustainable Development</td>
<td>Transnational Corporation</td>
<td>Vagueness in Spatial Data</td>
</tr>
<tr>
<td>Sustainable Development</td>
<td>Transnationalism</td>
<td>Vance, James</td>
</tr>
<tr>
<td>Alternatives</td>
<td>Transportation Geography</td>
<td>Varenius</td>
</tr>
<tr>
<td>Sustainable Fisheries</td>
<td>Trap Streets</td>
<td>Vectorization</td>
</tr>
<tr>
<td>Sustainable Forestry</td>
<td>Travel Writing, Geography and Tree Farming</td>
<td>Vernacular Landscapes as Expressions of Environmental Ideas</td>
</tr>
<tr>
<td>Sustainable Production</td>
<td>Trewartha, Glenn</td>
<td>Text/Textuality</td>
</tr>
<tr>
<td>Symbolism and Place</td>
<td>Triangulated Irregular Network (TIN) Data Model</td>
<td></td>
</tr>
<tr>
<td>Symptoms and Effects of Climate Change</td>
<td>Troll, Carl</td>
<td></td>
</tr>
<tr>
<td>Taphonomy</td>
<td>Tropical Rain Forests. See Biome: Tropical Rain Forest Tsunami</td>
<td></td>
</tr>
<tr>
<td>Taylor, Griffith</td>
<td>Tsunami of 2004, Indian Ocean</td>
<td></td>
</tr>
<tr>
<td>Taylor, Peter</td>
<td>Tuan, Yi-Fu</td>
<td></td>
</tr>
<tr>
<td>Technological Change, Geography of</td>
<td>Turner, Billie Lee, II</td>
<td></td>
</tr>
<tr>
<td>Telecommunications and Geography</td>
<td>Typhoons. See Hurricanes, Physical Geography of Typography in Map Design</td>
<td></td>
</tr>
<tr>
<td>Teleconnections</td>
<td>Underdevelopment</td>
<td></td>
</tr>
<tr>
<td>Television and Geography</td>
<td>Uneven Development</td>
<td></td>
</tr>
<tr>
<td>Temperature Patterns</td>
<td>United Nations</td>
<td></td>
</tr>
<tr>
<td>Temporal GIS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Temporal Resolution</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Terrain Analysis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Territory</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Terrorism, Geography of</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Text/Textuality</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Via Campesina (International Farmers’ Movement)
Vidal de la Blache, Paul
Video Games, Geography and Viewshed Analysis
Virilio, Paul
Virtual and Immersive Environments
Virtual Geographies
Vision and Geography
Volcanic Eruptions as Risk and Hazard
Volcanoes
Voronoi Diagrams
Vulnerability, Risks, and Hazards
Waldseemüller, Martin
Walker, Richard
War, Geography of
Waste Incineration
Wastewater Management
Water Degradation
Water Management and Treatment
Water Needs
Water Pollution
Watershed Management
Watershed Yield
Water Supply Siting and Management
Watts, Michael
Wayfinding
Weather and Climate Controls
Weber, Alfred
Web Geoprocessing Workflows
Web Service Architectures for GIS
Wetlands
White, Gilbert
Whiteness
Whittlesey, Derwent
Wilderness
Wildfires: Risk and Hazard
Wilson, John
Wind
Wind Energy
Wind Erosion
Wine, Geography of
Wine Terroir
Wise Use Movement
Wittfogel, Karl
Wood, Denis
Woodfuel
Woodlots. See Forest Fragmentation
World Bank
World Cities
World Court
World Health Organization (WHO)
World Summit on Sustainable Development
World-Systems Theory
World Trade Organization (WTO)
Wright, Dawn
Wright, John Kirtland
Writing
Xeriscaping
Zelinsky, Wilbur
Zoning
The Commonwealth of Independent States (CIS) is an organization founded to address economic and security situations facing the 15 former republics of the Soviet Union after the collapse of that country in 1991. The organization is a loose coalition addressing economic, trade, and security issues and has several constituent bodies. Headquartered in Minsk, Belarus, it is an organization that throughout its short history has been a weak organization despite initial fears that it was a vehicle for the recentralization of power to the Russian Federation.

### Context and Genesis

The demise of the Union of Soviet Socialist Republics (USSR) in 1991 led the republics to confront the economic and political realities facing them collectively. After years of domination by the Soviet Union, the constituent republics had a high level of interdependence and varying levels of experience in governing their own affairs. The various republics had developed economic specializations predicated on the Soviet Union’s series of Five Year Plans addressing economic development and production quotas.

This led to intrarepublic dependence on goods from newly independent countries that were formerly produced under the auspices of a single state. Formerly domestic trade had become international, necessitating the development of new agreements and bodies to govern disputes.

The organization was formed on December 8, 1991, by Belarus, Russia, and Ukraine through the Minsk Agreement. The group expanded to encompass another nine former Soviet republics with the Alma-Ata (now Almaty) Protocol on December 21, 1991. (Turkmenistan gave up full membership, opting for associate member status in 2005; see Table 1.)

The Charter of the CIS called for developing close ties and resolving disputes in several arenas, including the political, military, economic, social, and cultural. Article 2 of the Charter focuses on the settling of conflicts between Commonwealth states, while Article 17 requires members to resolve disagreements through negotiations and discussion or alternative procedures. The Charter allows for appeal to the Council of Heads of State if attempts to resolve the dispute fail.

### Military Cooperation

Originally, military cooperation was to run through the CIS Joint Armed Forces High Command (formed in 1992). The structure was reformed in 1993 as the Coordinating Staff. The divergent geopolitical goals of the member states have led to problems. Key issues include the drive by the governments of Ukraine and Georgia to align their countries with security organizations such as the North Atlantic Treaty Organization (NATO) over the question of objections by the Russian Federation and Russian actions in the breakaway Georgian provinces of Abkhazia and South Ossetia.

The Collective Security Treaty Organization (CSTO) was formed out of the CIS effort. The treaty was signed in 1992 by Armenia, Kazakhstan, Kyrgyzstan, the Russian Federation, Tajikistan, and Uzbekistan. Azerbaijan, Georgia, and Belarus joined in 1993, and the treaty came into force in 1994. The treaty was to last for 5 years initially. Three countries—Azerbaijan, Georgia, and Uzbekistan—declined to extend the treaty. The three states were joined by Ukraine and

### Table 1  Membership in the Commonwealth of Independent States

<table>
<thead>
<tr>
<th>Current Full Members (Year of Membership)</th>
<th>Associate Members (Year of Membership)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Azerbaijan (1993)</td>
<td></td>
</tr>
<tr>
<td>Belarus (1991)</td>
<td></td>
</tr>
<tr>
<td>Georgia (1993)</td>
<td></td>
</tr>
<tr>
<td>Kazakhstan (1991)</td>
<td></td>
</tr>
<tr>
<td>Kyrgyzstan (1991)</td>
<td></td>
</tr>
<tr>
<td>Moldova (1994)</td>
<td></td>
</tr>
<tr>
<td>Russian Federation (1991)</td>
<td></td>
</tr>
<tr>
<td>Tajikistan (1991)</td>
<td></td>
</tr>
<tr>
<td>Ukraine (1991)</td>
<td></td>
</tr>
<tr>
<td>Uzbekistan (2000)</td>
<td></td>
</tr>
</tbody>
</table>

Source: Author.
Moldova in withdrawing from the Commonwealth’s collective Security and Economic Cooperation Pact. Georgia, Moldova, and Ukraine withdrew to reduce dependency on Russian supplied energy, while the others sought to find alternative routes for oil and natural gas pipelines. The departure of these five countries leaves the Russian Federation in a dominant position vis-à-vis the other CIS members.

**Economic Affairs**

The intent of the CIS was to foster the continuation and enhancement of previously established trade linkages between member states. Under the centrally planned economy of the Soviet Union, the republics developed distinct specializations, leading to high levels of interdependence. The dissolution of the Soviet Union forced a reworking of economic relationships. Additionally, the republics were eager to develop trade linkages with other countries, but all remained cognizant of the markets they had and needed to maintain connection with.

Economic cooperation in the CIS is governed by the Economic Council and the Economic Court of the CIS. The latter is focused on ensuring that economic obligations taken on by member countries are fulfilled as specified by the CIS charter. The Economic Court comprises 24 judges. Sixteen are permanent judges, and the other eight are chief justices. Judgments by the court are only recommendations and, therefore, are not legally binding, thus weakening the effectiveness of the body as seen by several states’ refusal to submit to the court’s findings.

**Current Status and Prospects**

Observers note the relative weakness of the CIS and the continued dominance of the Russian Federation. Leaders of the CIS member states have noted the weak status of the organization, but it continues to serve a purpose for the members. Other groups have formed that are seen as challengers to Russian Federation dominance of the CIS. The GUAM Organization for Democracy and Economic Development was founded in 2001 by Georgia, Ukraine, Uzbekistan, Azerbaijan, and Moldova (Uzbekistan withdrew in 2005). The group, despite official denials, is seen as a move to counter Russian goals of continued regional dominance.

The CIS has been hampered by concerns by member states over the role of the organization in furthering Russian Federation interests. New alignments are being constructed as countries such as China, India, Japan, and the United States, along with European Union members, attempt to forge new relationships with CIS states. Concerns over the erosion of independence are also an important factor. The role of the CIS may change as leaders use it as a consultative organization where states can engage common issues in a political forum, but the possibility of the CIS becoming an organization with the strength of the European Union and NATO is seen to be unlikely by many experts.

*Darren Purcell*

**Further Readings**


**Communications Geography**

Communications geography is a nascent subdiscipline of human geography exploring the relationships between information and communication technologies (ICTs) and human conceptions of space and place. Historically, the unifying theme through much of the research was the movement of information and the physical geography of the networks making such a movement possible. More recent scholarship has focused on the socially constructed nature of technology, which engages the role of social processes in the creation and use of new technologies. This shift has resulted in research focused on issues such as the
global digital divide (gaps in access to ICTs that are along class and ethnic lines), the use of ICTs to achieve political and economic development goals, and the facilitation of community formation at multiple scales. Additionally, the subdiscipline has become an umbrella for cultural geographers researching the impacts of cultural products, such as the various media (television, film, advertisements, music, literary works, and performance arts such as dance). As a subdiscipline, the field is inherently interdisciplinary. Various literatures in political communication, communication, cultural studies, media studies, and Internet studies, as well as sociological work on ICTs, are incorporated into the understanding of the relationship between ICTs and space.

**Background**

Interest in communications and geography can be traced to the 1960s. Observers have noted that even then the rapid change in ICTs made it difficult to study the phenomena facilitated by new technologies. The earliest prominent texts were produced in the 1970s, with calls for geographers to engage communication practices and technologies more strongly. Much of this work was situated in the tradition of transportation geography and spatial analysis, with an emphasis on understanding the movement of data and the physical form of networks, both of which could be studied empirically. Such work was seen as supporting the progress-oriented, utopian visions of a technologically enhanced world in the 1970s. This strand of thought dovetailed with a view common among techno-optimists, such as Alvin Toffler, that space and distance were enemies to fulfilling human potential and that technology would usher in a new postindustrial era of peace and prosperity by conquering the tyranny of space. Despite this emphasis, some geographers were engaging social issues in terms of societal changes being facilitated by ICTs, asking about the resulting uneven social impacts. This cluster of work in the 1970s failed to take a strong foothold in geography, and engagement with ICTs dwindled by the end of decade.

The decline in engagement was shaped by a shift in the theoretical orientation of the discipline, namely, toward more humanist, Marxist, and behavioral approaches to geography, which criticized the spatial modeling approach to geography, which had dominated early work on ICTs and geography. What resulted was a shift in the discipline as a whole toward great engagement with concepts of representations of space (how places are portrayed), culture, and political economy. Unfortunately, research examining communication processes declined.

During the 1980s and early 1990s, the landscape of ICTs changed dramatically. The early Internet was being used by academics, and the foundations for the transfer of the network to the private sector were under way. Greater levels of digitization of analog data facilitated greater levels of transferability and economies of scale that transformed telecommunications. The breakup of the AT&T monopoly also ushered in an era of greater competition, resulting in new technologies being developed and offered to businesses and, later, individuals. Additionally, the decline in the cost of computing power and the exponential increases in processor speed and storage capacity facilitated the conversion of many forms of communication to digital form. The greater computerization and digitization of many social processes, such as converting voice communication to digital data and money to digital forms, and efforts to attain a “paperless society” have all necessitated adaptation to ICTs, providing social scientists much to study. Today, the processes are further enhanced as new ICTs are adopted by more firms, agencies, and individuals and formerly face-to-face communication is substituted for by social networking sites and text messages. Additionally, the overall diffusion of these computing networks, while uneven and not totally global in scope, reflects the fastest adoption of a communications technology the world has seen.

Research on the actual geography of ICT networks, the absolute and relative accessibility of the networks, and the impacts these have on the interactions between people in different places is still important, and these are more easily explored with advances in geographic information systems designed to facilitate network analysis. Another theme from the 1970s, the perception of ICTs and their use to facilitate strategic goals such as political or economic development, is still important across a range of subdisciplines. Communications
geography encompasses research concerns within the entire range of communication processes and technologies used to communicate ideas and data. These processes include older, dominant media such as radio, television, and film, along with existing print media that still have an influence on perceptions across the globe, despite the merging of their forms of content onto electronic platforms. The diffusion of fiber optic networks and 24/7 communication technologies, along with the exponential increase in computing power and speed, has created new opportunities for geographic analysis of ICTs and their impacts on societies.

**Dominant Themes in Communications Geography**

**Spatial Science and Modeling**

Communications geography has its roots in the thinking of the spatial science tradition and owes much to formulations related to transportation geography. Early efforts to engage the role of communications technology and human geography emphasized the manner in which such technologies were used to decrease the friction of distance, as all transportation advances have done. Instead of moving goods, the emphasis was on the movement of ideas and culture. The concept of time-space convergence was central to understanding the spatial impact of ICTs, the effect new forms of transportation or communication technologies had on the ability to move goods, people, or ideas in terms of communication. Other authors invoked the term *time-space compression* but linked it more closely with the changing geography of capitalism as it developed over time. Research following this tradition can be seen today, examining issues of accessibility to networks, the role of accessibility in facilitating economic growth, and personal usage, among other themes.

**Geography Is Dead, or Is It?**

A cover of *The Economist* magazine in 1995 proclaimed the “Death of Distance” thanks to the telecommunications revolution, with greater diffusion access to ICTs. Others went so far as to provocatively reference the end of geography in article titles, as ICTs were said to eliminate spatial barriers. Many geographers and other social scientists have expended (and continue to do so) great intellectual effort to rebut this position, arguing that if anything, geography is more important thanks to the ability to connect places with greater speed and data capacity. ICTs connect people in disparate places, thus there must be a reason to connect. The characteristics of the peoples and places being connected are vital to understanding the spatial patterns of usage. One example can be found in the literature exploring globalization and the use of ICTs to offshore production processes and information processing to places with lower labor costs. Other work focuses on the mapping of cyberspace and other communication and media flows, demonstrating that despite claims of universality and the ability to facilitate footloose movement, flows of power, money, and information follow earlier established patterns, with only incremental changes occurring.

**The City**

Early scholars working in the early utopian-influenced era considered the impact ICTs had on global urban structures, going so far as to question why cities such as New York would continue to exist in the decentralized world envisioned by various futurists, where telecommuting would be the norm and the centralizing impacts of the city would be waning. The present sees urban areas swelling across the globe but being reworked through the impacts of ICTs. Contrary to the futurists, the city and the multitude of daily interactions at the local and global scales have been intensified as people seek greater speed and mediation of their interactions. Cities that are well connected through infrastructure and information flows have prospered and continue to dominate the world system. The use of ICTs to aid in the governing of cities, the reworking of the urban form, and the economic prospects of cities in the future are topics often engaged by scholars. Much of this work has examined the local conditions for the increased production of knowledge, a resource vital in the competitive environment today.

**Cyberspace and Virtual Space**

The popularization of the Internet and the use of the term *cyberspace* led to the use of many
geographical metaphors to describe the “space” of interactions that were occurring online. Early Internet browsers invoked the terms explorer and navigator, implying the conceptualization of the technology as a place to be discovered by the individual, and books such as Virtual Community: Homesteading on the Electronic Frontier certainly evoked the idea of a place facilitating interaction. Chat rooms became spaces of meeting and exchange, and Web portals became focused on generating traffic. The technological hype about the possibilities of the Internet mirrored that of the early 1970s futurists (see, in particular, Wired magazine), and in similar fashion, geographers responded to the changes the new ICTs were facilitating, but in greater numbers and with a larger variety of methods. Studies focused on networks engaged the material connections that needed to exist to support the virtual worlds used by Internet browsers. This point was of particular interest as the Internet’s origin as a communications network for the military meant that the data traffic flows mirrored already existing geographic, namely, urban, hierarchies.

The idea of mapping cyberspace reflects a Cartesian view of geography, where to understand the phenomenon of the Internet, it must be mapped. The challenges of mapping a network that was growing exponentially were undertaken by several, but most notably in the book Atlas of Cyberspace. Further work in geovisualization of data has aided such efforts.

Theoretical engagement about what is a virtual space engages our understanding of the term virtual and notes that many things taken as real can be classified as virtual, and vice versa. The virtuality of the phenomena does not matter, as individuals still choose to act as if they were real, bringing into question the notion of the virtual world as constructed separately from the material or real world. As the Internet has attracted large numbers of people, it has come to form a virtual public space of debate, dialogue, and many other forms of interaction.

Interaction and Social Consequences

The social impacts of the use of ICTs has been very important to consider. A burgeoning body of work has engaged the idea of generational differences in the use of the Internet and its construction as a threat to children in various news media and policy outlets. Others have explored the role of ICTs in areas of conflict and rapid cultural change, such as the Middle East, and the various restrictions placed on the use of the Internet and other ICTs. In general, one research trend has been the study of groups in various societies and their adoption of ICTs, how they differ and how they are deployed in ways that transform societies through their use.

Economic Changes

Much of geography’s engagement with ICTs has concerned economic geography, in particular the issue of globalization of services. Service industries lend themselves to taking particular advantage of ICTs and the new geographies they fostered. Early studies noted the impact of ICTs on legal services, offshore finance, and the vertical disintegration of many firms as they concentrated management and strategic operations in large cities, where telecommunications infrastructure was available, while back office functions such as forms and data processing were offshored to sites with infrastructure but cheaper labor than in the core regions. Additional contributions have explored the level to which various societies have become “information societies.” The term is debated, as observers point out that while services (industries usually seen as information processing) are expanding across the globe, the increased use of ICTs across all industrial pursuits, including primary and secondary economic activities (such as farming, fishing, and the processing of their products), in addition to manufacturing, has changed the nature of such work. Observers argue that all work has been information rich in terms of tacit knowledge, but with ICTs, the ability to track and replicate such information is enhanced, making the information age not a new era but one of greater emphasis on the informational nature of all work.

Additionally, the role of place in fostering the Internet industry has been examined within the larger theme of economic innovation and development, critically engaging the idea of fostering future Silicon Valleys and Silicon Glens across the globe.
The 1990s and early 2000s witnessed an expansion of work engaging the power of various media to frame places for consumption and as a political power resource. Early work was critical in that it engaged the idea of using the power of the media to craft place images that served as a resource or a dominant frame by which to understand places. Later work, a subset of critical geopolitics, engaged the production of what is termed popular geopolitics, the representations of the place that reflect popularly held geopolitical views. These representations of places are considered geopolitical acts themselves, as they structure what the populace may know about areas of strategic concern to policymakers and elites. A variety of media have been explored in conjunction with the role of elite-produced texts that inform mass publics about the priorities and concerns of various states, such as news magazines, movies, and comic books. Much of the research has concluded that the subtle power of the media to shape geopolitical worldviews is quite real and not questioned by those who consume such media products. Newer work has engaged the issue of communication platforms as a space of engagement for political struggle.

The political possibilities of ICTs have been discussed since the 1970s by utopian thinkers, but the realities of who controlled the media and ICTs became more apparent in the 1980s. Observers in the 1990s noted that the introduction of ICTs only seemed to concentrate power in the core regions of the world and worked against the empowerment of the masses to facilitate greater levels of democracy and political participation. Concerns over the increasing surveillance and self-surveillant nature of ICTs were expressed, in particular after the September 11, 2001, terrorist attacks in the United States. Much of this work can be found in writings focused on the development of closed-circuit television systems in cities. Other authors have conceived of the Internet and its spaces for discussion as a new form of Habermas’s public sphere, though critics argued that the Internet’s capabilities to facilitate communication across groups is limited as the categories of the material world such as class, race, gender, and ideology are present in the virtual world, thus guiding users to likeminded sites and, thus, only reinforcing preconceived notions.

The subdiscipline of communications geography continues to grow in terms of status, the types of research conducted, and the numbers claiming an interest in related research. Methodologically, the subdiscipline is very heterogeneous, with researchers employing network analysis, quantitative measures via geographic information systems, and qualitative methods such as textual analysis and interview techniques. This mirrors the broad variety of methodologies seen in other disciplines that engage communications and media questions. The Association of American Geographers hosts the Communications Geography Specialty Group, which was officially recognized as a specialty group in 2003. Other geography associations are interested in communications, such as the International Geographical Union (IGU), which also has a commission on the Geography of the Information Society. Other academic associations’ conferences also see an increasing presence of geographers as members and in the number of sessions devoted to spaces, place, globalization, and ICTs.

Darren Purcell

See also Cyberspace; Media and Geography; Telecommunications and Geography; Television and Geography

Further Readings

Communism, the political and economic regime characterized by a command economy and single-party rule dominated by the members of the Communist Party, wrought significant effects on the cultural, urban, economic, and physical geography of communist-ruled states in both Europe and Asia. While there are differences between the way communist goals were pursued in the Soviet Union after World War I and in Eastern Europe and Asia in the second half of the 20th century, there are a number of commonalities. Because of the control exercised by the Party, communist countries were able to be single-minded in their pursuit of goals, as exemplified by the proliferation of Five Year Plans, especially in the areas of industrialization, agricultural restructuring, and urban morphology.

**Industrialization**

One of the most striking aspects of communist rule, especially in the Soviet Union and Eastern Europe, was the implementation of quick and massive industrialization. While there had been some industrial development in Russia and Eastern Europe prior to World War II, notably in Poland and Czechoslovakia, Stalin implemented an aggressive program of rapid industrialization to raise the USSR to the level of the advanced capitalist countries within 10 years with specific focus on heavy industry and defense. Estimates of the annual growth in industrial production during the first two Five Year Plans (1928–1937) range from 10.5% to 18%, with dramatic increases in power production, blast furnace capacity, steel production, capital stocks, and factories. After World War II, industry remained a priority, with the development of industrial complexes, especially east of the Ural Mountains and in Central Asia. These sites, such as the one in Ust Ilimsk, featured natural resource extraction, huge factories, and large cities built often ex nihilo. Rapid industrialization was also undertaken in the satellite states so that former agricultural countries such as Hungary and Poland witnessed the development of industrial centers where once there were only small towns.

In China, the government also undertook nationalization and rapid industrialization once communist control was established in 1949. During the 1950s, China increased production in the industrial sector with annual growth rates of 8% to 10%, though small-scale rural industrialization also received special attention. During the Great Leap Forward (1958–1961), for example, “backyard furnaces” were developed to smelt scrap iron, though this proved to be untenable. This period of experimentation resulted in reverses in industrial growth and negative effects on agricultural production as a result of the diversion of labor. However, the groundwork for more successful small-scale and rural development was laid, and the 1960s and 1970s witnessed the creation of rural factories and widespread electrification through hydroelectric power production. The controversial building of the Three Gorges Dam is the latest chapter in providing power for the country.

**Collectivization**

Communist states also pursued collectivization, effectively nationalizing agriculture into large...
cooperatives. In the USSR, Stalin implemented an ambitious program with two aims: to destroy the power of the kulaks (wealthy landowning peasants) and to increase yields to provide food for the growing urban workforce. The number of agricultural workers declined from 72 million in the mid 1920s to between 48 million and 53 million in 1939. Both state farms and collective farms existed, and strict production quotas were imposed. Peasants were allowed individual garden plots, which, because of their high productivity, were an important piece of the Soviet agricultural picture. The results of collectivization, especially during the first Five Year Plan and again in the period after World War II, were less than successful. While production of cotton and potatoes increased, livestock numbers dropped precipitously, and large-scale famines occurred in grain-producing areas such as the Ukraine in the early 1930s and again in the 1950s. Despite this, agricultural growth was on par with that in the rest of the world throughout the 1960s and 1970s. In the Soviet satellite states, collectivization was also imposed with varying degrees of intensity.

China in the early 1950s undertook gradual collectivization, introducing cooperative use of tools and then the grouping of land purchased from peasants, with initially positive impacts on grain production. During the Great Leap Forward, collectivization was pushed forward with the aim of organizing social life in the countryside. Large-scale communes consisting of up to 5,000 families were created, and private plots were abolished. In some places, family life was severely controlled with the establishment of single-sex dormitories and children’s nurseries. Dissatisfaction, combined with weather-induced poor harvests and subsequent mass famines with estimates of 17 million to 30 million dead, led to a return to the more successful cooperative model by the 1960s. North Korean collectivization also followed the commune model, organizing production as well as nonfarm services such as education and health care. In the 1990s, North Korea was plagued by widespread famine, resulting in death, chronic malnutrition, and political defection. In Vietnam, collectivization was undertaken in the North with allocations for a small percentage of collective land to be held privately. After 1975, attempts were made to collectivize South Vietnam, with little progress, and the effort was eventually abandoned.

**Communist Cities**

Communist rule brought changes to urban morphology—namely, in the areas of housing, architectural style, and public space. While urbanization was actively pursued in Europe, in China urbanization was more controlled, resulting in special emphasis on the development of rural areas and distinct periods of counterurbanization (e.g., the Cultural Revolution in 1966–1969). Cambodia under the Khmer Rouge (1975–1979) presents an extreme example of counterurbanization.

**Architecture, Ideology, and Public Space**

In the period immediately following the Russian Revolution, architecture was viewed as an ally in achieving utopian goals; modernism and the Russian avant-garde were embraced throughout the 1920s. However, formalism came under suspicion in the 1930s and was abandoned in favor of the neoclassical forms of socialist realism. Characterized by buildings on a monumental scale and the use of classical columns and sculptural elements and reliefs featuring the proletariat, socialist realism predominated in the USSR and the Eastern Bloc from the 1930s through the mid 1950s. Eastern Bloc architecture has since been characterized by modernist industrial building techniques represented by the bloc flat.

Public space in communist countries was uniquely marked. Streets, squares, and cities were viewed as potent canvases on which to project socialist ideology, resulting in the wholesale renaming of places after the heroes (e.g., Lenin, Ho Chi Minh), the martyrs (e.g., the Rosenbergs), and the ideals (e.g., people’s republics) of the movement. Cities and towns were inscribed with posters and monuments conveying didactic messages about the regime, depicting the heroic proletariat and women as mothers and workers and propagating the “cult of personality.” Notable was the scale of public monuments, such as the Ernst Thälmann monument in Berlin (which was
COMMUNISM AND GEOGRAPHY

13 m [meters] high), the 25-m-high Stalin statue in Budapest, and the 170-m-high Juche Tower in Pyongyang. Boulevards and squares were crafted or coopted as sites of state-sponsored events, for example, the May Day parades featuring military hardware, workers, and youth (see photo). Building types specific to communist cities also marked the landscape, including trade union headquarters, so-called people’s palaces, pioneer camps, and culture houses, which were found in cities and small villages and offered programs such as folk dancing and youth activities in support of the ideological goals of the regime.

**Housing**

In Europe, the concomitant urbanization that resulted from rapid industrialization, coupled with the wartime destruction of cities, meant that housing presented an ongoing challenge to communist states. In the aftermath of the Russian revolution and later after World War II, private property was confiscated, and large villas and apartments were split into smaller units. Industrial techniques including the production of ferroconcrete slabs were enlisted in the 1950s to alleviate the ongoing housing shortages, resulting in the construction of multistory housing estates, which are a characteristic feature of socialist cities. Known in Russian as *microraion*, these estates were organized to house hundreds of thousands of people. As they were often built on the urban fringes, they were linked to efficient mass transportation and provided on-site services such as nurseries, shops, cultural centers, and parks.

Michelle Marie Metro-Roland

May 1, 1987: In Moscow, a banner showing Friedrich Engels (center), Vladimir Lenin (right), and Karl Marx (left) hangs over Red Square crowd during May Day celebrations.

Source: Time & Life Pictures/Getty Images.
COMMUNITY-BASED CONSERVATION

Community-based conservation is one term of many describing the direct involvement of local people and organizations in decisions and activities associated with locally based environmental management. Its introduction marks a shift that includes humans in ecosystems and a shift from management by experts toward management that includes local perspectives and interests. Other related concepts include comanagement, ecosystem management, civic science, grassroots environmental management, and public ecology. Its rationale is rooted in the subsidiary principle that decisions should be taken at the level closest to where effects are most likely to be felt. This implies that the people most affected should have a direct influence on the decisions that are taken. Proponents of community-based conservation typically also promote ecosystem management, a management approach that attempts to consider the environment holistically and explicitly works to restore and sustain healthy ecosystems, including their multiple functions and values.

Many benefits have been attributed to community-based conservation. Governments and private proponents may find that they can be more efficient in the long term if their plans are locally discussed prior to implementation and they involve local people directly in management efforts. Others have suggested that community-based conservation may also democratize scientific knowledge by linking experts and stakeholders in planning social, economic, and environmental improvements.

Furthermore, it has been suggested that by working together, people can accomplish more than as individuals or as organizations working in isolation. Additionally, involving those affected will provide for a broader range of values and inputs into the decision-making process. Local people may have specific knowledge about their local environment that can help tailor conservation to local conditions. Involving local people may increase the likelihood that decisions and practices are accepted locally, while encouraging compliance through peer pressure and by setting good examples locally. Involvement of local people in community-based conservation activities may also help build and enhance local skills, interests, and capacities that will continue to serve these communities in other areas. Finally, community-based organizations may be more nimble than larger institutions and so may be able to harness opportunities within smaller areas efficiently and effectively.

Activities associated with community-based conservation are wide-ranging. They may include undertaking environmental cleanups, bird counts, or habitat restorations that may or may not involve ongoing commitments. They may involve activities that supplement government functions related to environmental conservation, including public education, monitoring, and construction of facilities. They may model environmentally friendly behaviors such as conducting creative experiments or establishing land protection programs. Increasingly, the term also refers to activities that involve partnerships among government, private sector, and citizen organizations where...
groups share information, labor, money, and even decision-making authority, including advisory or comanagement committees.

Evaluating the success of community-based conservation is challenging. Many assessors focus on procedures whereby criteria such as holding a broadly shared vision or establishing an open, accessible, and transparent decision-making process are important. Others suggest that both socioeconomic and procedural outcomes must be addressed. Criteria focused on outcomes might assess changes in employment, reductions in expenditures, or improvements in habitat conditions or water quality. The challenge, however, lies not only in determining appropriate criteria and acquiring supporting data but also in being able to attribute ecological or social changes to specific conservation activities.

Notwithstanding possible benefits, several concerns have also been raised about the efficacy and appropriateness of community-based conservation. First, there is concern that promoting active citizenship at the community level encourages governments to abrogate their responsibilities, so that tasks once delivered by the welfare state are taken up by the local community. Paradoxically, communities may find themselves disempowered if resources once provided by central levels of government are reduced while expectations of what communities can achieve locally mount. Second, community-based conservation requires legitimacy both locally and extralocally so that participants will have an incentive to contribute.
However, concerns expressed by scientists and industry and government representatives frequently revolve around whether local residents have the capacity or the expertise to engage in conservation efforts, thus reducing the incentive for local residents to take part. Third, questions are also raised about how local people are selected and to what degree local individuals may be representative of the diversity of interests at the community level. Local elites may use the rhetoric of fairness and equity to mask the pursuit of self-interest and entrenchment of power at the local level. Local communities are not necessarily more democratic than centralized authorities as power relations within communities may exclude some social groups, stifle debate, and/or generate a form of paternalism that is locally generated.

These concerns also raise issues about the uneven capacity of communities to maintain or restore healthy ecosystems, local livelihoods, and associated sociocultural values. Due to differences in circumstances, such as the economic base, natural resource endowments, proximity to urban locations, and demographic characteristics, some communities may have a relative advantage over others in providing environmental services. Furthermore, local groups may vary in the strength of their internal structure, the ability to obtain funding, their willingness to take risks, and their effectiveness in networking with other organizations and agencies. Thus, by emphasizing local programs, geared to local ecological and social conditions, community-based conservation may contribute to highly differentiated application of conservation practices.

These concerns have now given way to understanding that no level of governance—community or nation-state—can manage ecosystems for conservation in isolation from other levels. Instead, community-based conservation now also focuses on ways to ensure that decision-making processes are legitimate, accountable, and inclusive of multiple stakeholders and interests. To meet these conditions, all the actors in conservation should establish strong linkages across spatial and temporal scales, and authority should be spread across multiple institutions to sustain and improve conservation practices.

Maureen G. Reed

See also Community-Based Environmental Planning; Community-Based Natural Resource Management; Indigenous Environmental Knowledge

Further Readings

particular knowledge or social connections to places have also been given strong recognition in planning processes.

Another important aspect of the ancestry of CBEP is the well-chronicled failings of the so-called top-down, or rational-comprehensive approach to planning, which, it has been argued, is insensitive to the particularities of locality and, thus, poorly adapted to human and ecological diversity across landscapes. In addition, top-down planning has been deemed authoritarian, implementing its prescriptions with insufficient sensitivity to and regard for local values and interests. These problems, as J. C. Scott has powerfully argued in *Seeing Like a State*, result in the failure of top-down planning to understand and use local, *indigenous* knowledge, which is crucial to designing and implementing effective interventions.

**Advantages**

Both the scholarly literature and practice in diverse circumstances across the world provide evidence that CBEP has a number of advantages. First, CBEP is more sensitive to the local characteristics of a given environmental problem. Essentially, place-based communities are thought to be more intimately familiar with environmental problems and solutions and can incorporate local knowledge to better inform the planning process, which in turn increases awareness of environmental issues.

Second, the planning outcomes that result from CBEP enjoy higher levels of legitimacy because they involve local residents and interested parties and promote “ownership” of problems. A common limitation of state-directed environmental planning has been a lack of legitimacy, which has hindered acceptance of planning outcomes. This is, indeed, part of the raison d’être of CBEP: By facilitating the participation of local, place-based actors, CBEP has the potential to avoid community resistance to plan implementation. As a result, CBEP has been seen as a more effective and more efficient mode of planning.

A further advantage of CBEP concerns the fundamental role of community participation in plan development. Relating to Jürgen Habermas’s concept of communicative action, citizen involvement in planning processes is perceived to have a democratic benefit in its own right. This benefit stems from the importance of deliberation in the form of questioning, rationalizing, and comparing different perspectives in the planning process. On this basis, even if the planning outcome was identical to a state-dominated process, the social learning that occurs in developing an environmental plan has intrinsic value and is central to a healthy democratic process.

**Problematic Aspects**

There are, perhaps inevitably, some problematic aspects to CBEP. The first of these relates to the character, role, and functionality of “community.” Common conceptions of community imply a homogeneous, spatially fixed social group that shares a consciousness of being a “community” and that is characterized by social consensus and solidarity. Such a view has been subject to critique, particularly by those concerned with understanding how “difference”—gender, ethnicity, class, age, and other forms of social identity—divide so-called communities. Second, empirical reports of the practice of CBEP suggest that this mode of planning has tended to empower powerful and articulate actors and marginalize others. These reports directly undermine one of the perceived advantages of CBEP—its democratic effects.

A third problem with CBEP relates to the capacity of communities to plan and implement strategies for the management of environmental assets and problems—given the technical complexity, time frames, and scales at which some environmental problems are manifest. Fourth, one of the promises of CBEP is to enable the integration of local, experiential (indigenous) knowledge with scientific knowledge. Questions have been raised as to whether community-based approaches can harness indigenous knowledge while retaining technical fidelity in their response to environmental problems. Others have questioned whether different and competing knowledges can be reconciled in a single planning process.

Finally, CBEP has been criticized for being maladapted to the scale of contemporary environmental issues. There are multiple issues here,
including the following: the potential for parochial thinking to dominate, the possibility that strategies that seem appropriate at local scales can have harmful effects at wider scales, and the possibility that environmental planning is focused on the symptoms rather than the drivers of environmental problems.

**Potential Applications**

These complications suggest that CBEP cannot be universally applied; its utility is limited to particular circumstances. The first is when the “community” involved has established a degree of cohesion, a shared history, and mechanisms for resolving differences. The second is when CBEP receives institutional support from the wider governance system, including adequate resources to facilitate deliberation. Finally, CBEP is more likely to work at smaller spatial scales and when there are finite and discrete planning concerns to be addressed.

*Thomas G. Measham and Marcus B. Lane*

See also Community-Based Conservation; Community-Based Natural Resource Management; Community Forestry; Environmental Planning; Participatory Planning

**Further Readings**


Community-based natural resource management (CBNRM) has been characterized as both a practice, the management of local resources by a geographically limited community, and a process, the shifting or devolution of power and authority from bureaucrats working within national or state governments to more local and often indigenous institutions and peoples. Though theoretically feasible for any number of natural resources, in practice, community-based management has been almost exclusively limited to flow resources such as forests, fish, and game, which, as common property, have tended to suffer from overexploitation. CBNRM has received increasing attention of late in geography and related disciplines, which likely stems from the fact that it stands in marked contrast to the still dominant distant-centered model of resource management used throughout the world and is perceived as superior to it, especially with respect to ensuring long-term resource sustainability.

**CBNRM as a Practice**

In practice, CBNRM is probably best recognized in innumerable community forests that exist throughout the world, some of which, as with Italy’s Magnifica Comunita di Fiemme, have existed for many centuries. Fish and game have also commonly been subject to community scale management, as exemplified by Zimbabwe’s Communal Areas Management Programme for Indigenous Resources (CAMPFIRE), through which small-scale rural district councils manage lucrative sport hunting of elephant and other wildlife and the harvest of lesser local resources such as crocodile eggs, timber, and river sand. The defining feature of these and other cases of CBNRM is that local landholders own, or at least gain access and use rights to, valued local resources, which they manage in their interests. This feature makes CBNRM distinct from comanagement (whereby communities share authority with governments), citizen-based environmental
management (whereby governments manage resources based on citizen input or, as in the case of California’s “Ballot Box Planning” approach, voter dictates), and even civic environmentalism (whereby community groups assume responsibility for the cleanup of local environmental problems).

Beyond this core criterion of ownership or access rights, reviewers have regularly identified, or at least promoted, several other defining characteristics of CBNRM, such as the use of participatory models of decision making within the community; the promotion of social equity and justice, especially for traditionally disadvantaged community members; the inclusion of traditional knowledge and values in the management of community resources; and the reconciling of the typically antithetical objectives of economic development and resource conservation.

Unfortunately, real-world cases of CBNRM have often been challenged to achieve these laudable goals, and indeed, many have never sought to. Research on community forestry as practiced around the world offers insight. While not capturing all instances, reviewers have come to identify three distinct forms:

1. Community forestry as ecological forestry, whereby managers seek to halt deforestation and improve ecological conditions
2. Community forestry as social forestry, whereby timber and nontimber resources are promoted as a basis for improved subsistence or modest commercial enterprise among poorer strata
3. Community forestry as essentially small-scale industrial forestry, whereby the unapologetic avowed goal is economic gain

In other words, CBNRM is commonly characterized in geographic and other disciplinary scholarship in idealized terms that may not match its real-world practice.

CBNRM as a Process

For many reviewers, the significant aspect of CBNRM is not the practice of resource management by communities but rather the process by which these communities come to secure management authority over local resources. This focus of interest reflects not only the phenomenon’s still nascent status but also the considerable challenge to devolve power fully from distant-centered bureaucrats to community institutions in many parts of the world. Indeed, even well-known cases of CBNRM often reflect incomplete forms of devolution. In the case of CAMPFIRE, for example, the Central Zimbabwian government must approve the proposed harvesting rates. The same is true for British Columbia’s (Canada) many community forests, which must negotiate with provincial officials to establish their annual allowable cut.

Though incomplete in too many instances, devolution continues to be promoted around the world for a variety of related reasons. For example, in countries and subregions where colonialism resulted in the expropriation of traditional landowners’ resource use rights, CBNRM has been advanced as a necessary step to redress past injustices. These cases challenge the view of CBNRM as a novel phenomenon, so asserted Stephen Kellert and colleagues in a 2000 study; rather than being new, the authors argued, CBNRM can be viewed as a modern attempt to revive traditional, cultural, and institutional mechanisms for managing and conserving natural resources. This latter aspect, the expectation that resources will be better conserved, constitutes another key driver of devolution. In these same colonial contexts, observations of overexploitation of formerly closed- but now open-access resources eventually led governments to impose Western-style regulatory systems (based on external rules and enforcement) in an attempt to stop this tragic outcome. Given the limited success of these systems, not just in former colonial contexts but also in places such as Canada’s east coast, where cod stocks were harvested to exhaustion in the 1980s, CBNRM has been promoted as a practical means of addressing the problem of common resource overexploitation.

CBNRM as Sustainable Management?

It is not enough for advocates of CBNRM to simply discredit the capacity of central governments to sustainably manage flow resources that are harvested by multiple resource users, typically at great distances from the place of management;
they must also identify the basis for expecting that CBNRM can do better. Fortunately for them, there is ample logic and considerable evidence that supports this expectation. Drawing on the widely embraced principles of adaptive harvest management, which calls on managers of flow resources, such as a fishery, to be prepared to alter a management decision if it is later found to be ill conceived, CBNRM advocates contend that this is only achievable, or at least best achieved, if the manager is there, on site, to see how a system responds to a management action. While it may be reasonable to accept that community-scale managers can respond faster with site-specific solutions than can distant managers, what ensures that they will want to do so, especially if the solution entails reducing the community’s harvest? To address this concern, CBNRM advocates commonly invoke two arguments. The first, which Bonnie McCay and Svein Jentoff labeled the feedback effect in a 1996 study, posits that local management, unlike distance-centered management, directly subjects decision makers to the repercussions of their decisions and thereby promotes more conservative decisions. The second, which Robert Paehlke labeled the bioregionalist promise in a 2001 study, posits that communities of people who have lived in a particular place for generations and thereby gained a “sense of place and belonging” are expected not only have superior local ecological knowledge but, more important, a vested interest in ensuring the survival of that place.

Innumerable historic and present-day cases of CBNRM support this logic; however, there are other cases that challenge it. Based on emergent research in geography and related disciplines, the determinants of “successful” CBNRM, whether it be defined in economic, social, or environmental terms, are gradually being revealed. In the case of environmentally oriented CBNRM, it is increasingly becoming clear that if communities can secure sufficient ownership and authority over a local resource such that resource users benefit from its sound management, especially in financial terms, stewardship of the resource will likely follow. This likelihood has afforded CBNRM increasing attention and promotion.

Ben Bradshaw

See also Adaptive Harvest Management; Common Property Resource Management; Community-Based Conservation; Community Forestry; Environmental Management; Indigenous Environmental Practices; Indigenous Forestry; Sustainable Forestry

Further Readings


Community Forestry

Community forestry is collective management of forests by local residents, often in collaboration with diverse interests that have competing claims on forest resources, to enhance community well-being and improve forest health. It entails reconfiguring relationships between communities, the
state, and other groups that claim rights to ownership and the benefits of forest management. The goal is to rectify the negative environmental and social impacts of industrial-scale forestry, which has often had significant backing from the state. Although community forestry is intended to build the capacity of communities to steward natural resources and have a voice in their own affairs, it faces challenges in enabling broad and meaningful participation, defining community, and sustainable forest management.

Participation and Building Community Capacities

Colonial administration of forestlands in developing countries led to the breakdown of traditional local natural resource management institutions, degradation of forest ecosystems, and impoverishment of local peoples. While conservation and poverty alleviation programs were implemented after independence from colonial rule in the 1950s and 1960s, by the 1970s, it was clear that they were failing. Professional forester, conservationists, and development practitioners began questioning state control of forestlands. They advocated reorienting forestry to management for multiple forest values, and to ensure that this new management benefited local peoples, they called for participation by community members in defining how forests should be understood and managed. A decade later, similar reasoning was applied in the United States, where communities experience limited access to forest resources, restricted opportunities for participation in forest decision making, and, at that time, heightened conflict over logging and endangered species protection. In hundreds of communities across the United States, local residents, scholars, and land management professionals have come together seeking innovative ways to collaborate on forest management.

Community members often need new skills to navigate the current global political economy, and capacity building is a central component of community forestry. Nicholas Menzies observes that community members need capacities in five areas: (1) building local institutions for setting priorities and mediating conflicts, (2) representing their interests in wider social and political arenas, (3) managing resources sustainably, (4) participating as informed actors in markets, and (5) building community assets with benefits derived from managing the resource.

Building local capacities requires working with community members as equal partners and reshaping relationships of power so that local people are positioned to make decisions about policy formulation and implementation. Participatory research developed hand in hand with community forestry as a means of increasing community capacities. When done well, it facilitates mutual learning among professionals and community members in ways that can, under the right conditions, enhance community members’ ability to apply research results to improve their situation. Yet people in positions of authority—whether local, regional, or national—are often reluctant to relinquish power. Communities, state agencies, and other interested parties worldwide struggle to negotiate the equitable distribution of rights, powers, and responsibilities.

Because of unique combinations of historical, economic, political, social, cultural, and environmental factors, these negotiations lead to unique tenure and sociopolitical arrangements in every place. Some examples of the multiple forms of community forestry include joint forest management in India, in which state forest agencies and local communities share rights and responsibilities for forest management; communal ownership of land in Guatemala; and collaboration among government agencies, community groups, environmental organizations, and lumber companies on the management of U.S. public lands. Whatever the case, the central concerns common to all community forestry efforts are clearly defining rights of access to forest resources, specifying protocols for participation in decision making, and enforcing rules to ensure that forest use is sustainable so that it is economically efficient, socially equitable, and environmentally sound. To community members, this often means working to maintain a particular way of life.

Central Issues in Community Forestry

Many difficult issues arise in establishing rights and protocols. A major issue is the enabling of participation itself. The institutionalization of
participatory development and research as embodied by requirements for local participation in development projects sponsored by international development organizations, for example, the World Bank, has led some scholars to describe participation as a “new tyranny.” They are concerned about participation being done only to fulfill a requirement of the sponsoring agency or because it is fashionable. Under such circumstances, community members are not given real power to influence forest management or economic development, and participation becomes nothing more than an empty promise.

But an empty promise for whom? Just who is “the community” in the first place? “Community” defies definition because people associate with one another in many different ways and for many reasons. There are diverse interests within communities that may conflict, just as national and local interests often do. How a community is conceived accords strategic political advantages to some groups while effectively excluding others. Racial and ethnic minorities, women, workers, and other marginalized groups find that their interests are often not well represented or addressed by community forestry.

Even when a broadly inclusive participatory process is achieved, a question remains as to what extent communities can manage their resources sustainably. While local communities may be good land managers when ecological, economic, and social factors are favorably aligned, community forestry is no guarantee of high-quality environmental stewardship.

Evidence suggests that secure rights to access forest resources and participate in decision making as well as strong local-level resource management institutions are necessary but insufficient to help communities halt environmental degradation and improve the well-being of their members. The challenge lies in delineating rights and responsibilities among the many different interests involved in each case and creating conditions conducive to making choices about the use and management of forests that are equitable, environmentally sound, and adaptable to changing economic, social, political, and environmental conditions.

Carl Wilmsen

See also Common Property Resource Management; Community-Based Natural Resource Management; Participatory Rural Appraisal; Political Ecology; Social Forestry

Further Readings


The term commute has become shorthand for traveling some distance between one’s home and place of work on a regular basis. The term commuting is derived from commutation ticket, the U.S. term for a season ticket because the daily fare is commuted to a single payment. Commuting clearly implies the routine and repeated nature of a work-related journey. Hence, commuters are defined as those who travel to work (sometimes also to school) by mainly motorized transportation (private car or public transport), and non-commuters are those who work at home or walk or cycle to work.

Policymakers and researchers are very interested in collecting statistics on commuting behavior and studying commuting patterns and trends in and around metropolitan areas. The primary objective is to better understand the variations in commuting distances and times, modal usage and choices, and trip planning and chaining behavior among different groups. Usually, a multidimensional approach is followed, whereby economic, social, health-related, and spatial aspects come to the fore.
The main factor behind the increased demand for mobility in general and commuting in particular has been the geographical dispersion of economic activities, which in turn resulted from the effect of motorization and the related substantial increase in the use of the private car. It is true that most people have become more mobile and footloose and experience greater locational freedom, which in turn has led, in many modern, industrialized societies, to a clear trend of moving away from traditional older urban centers in favor of suburban and rural locations. The result is urban sprawl and suburban development. This separation of the place of work and the place of residence has led to a substantial increase in commuting distance and time. Other noticeable trends that also have an impact on commuting are the increasing number of households in which at least two family members work at distinct locations other than the place of residence, the increase in the number of solo drivers (the so-called single-occupant vehicles, or SOVs), and a higher average disposable income, resulting in a higher number of car owners.

Studies in both the United Kingdom and the United States have shown that the average one-way travel time for a commuter in 2000 was about 25 min. (minutes), which represented a 3-min. increase in travel time over the average in 1990. This increase in commuting distances and times, leading to longer peak hours, also brought into the open the issue of long-distance commuting and excess commuting. In general, long-distance or extreme commuting occurs when a daily journey to work takes more than 90 min. each way, and this phenomenon has contributed to what economists term the commuting paradox. The commuting paradox refers to the fact that people usually overestimate the value of the things they will obtain by commuting (e.g., an intrinsically or financially more rewarding job, additional welfare gained from a pleasant living environment, more material goods, lower rents for housing, and more prestige) and underestimate the benefit of what they are losing: social connections, hobbies, and/or health. One extreme was noted in 2006, when America’s Longest Commute award was given to a person who drove 372 miles (or about 4½ hours) roundtrip to and from work each day. Excess commuting, on the other hand, is defined as the additional journey-to-work travel represented by the difference between the average actual commute and the optimum (or minimum) possible average commute, given the spatial configuration of workplace and residential sites. This concept is interesting because it relates to the issue of defining and modeling what could be termed an ideal commute time and a relative desired commute amount. Obviously, a single answer is not possible here because of the huge heterogeneity in commuting behavior.

It is universally assumed that commuting (and by extension all travel) is considered a source of disutility—a burden or a hassle—given the conventional derived-demand justification for travel. In other words, commuting always has “its price,” and people seek to minimize its cost, subject to a number of constraints. This cost is the monetary cost of congestion and journey delay, calculated on the basis that time spent traveling is assumed to be nonproductive and wasteful to the economy. In economic terms, this translates time into a monetary value (i.e., the principle of value of time, or VOT) using the average wage of the commuter. There are also additional, unpredictable costs due to the unreliability of travel time, and of course, there is the (generalized and out-of-pocket) cost of travel itself. Studies in the United States have indicated that a typical household spends nearly 20% of its income on driving costs, which is more than it spends on food. Besides these private costs, there are the social costs of commuting, due to congestion, noise, and pollution of the environment.

Commuting not only takes time and generates costs but also has important physiological and psychological impacts. For one, there is the physical and mental burden of stress, health risks, fatigue, and combining work with family life linked with the nature of the trip to work. Delays caused by traffic volume (congestion), accidents, or bad weather conditions; the driving behavior of other road users (for car users); the unreliability and the lack of frequency of public transport services; the negative effects of crowding or unforeseen strikes (for public transport users); and the imbalance between work and family (especially when children are involved, who need to be dropped off, picked up, and cared for), all
contribute to increased travel stress levels. Moreover, when there is less control over these factors, commuting becomes more stressful. The results are well-known: high blood pressure, stiff neck, tiredness, self-reported tension, aggression, negative moods, increased sense of frustration, more frequent illness, reduced task performance, work absences, less job stability, and overall lower life satisfaction. Note that all aspects typical of less commuter-dependent jobs (such as jobs with flexible working schedules, part-time work, teleworking, and jobs closer to the home location) have a strong stress-reducing impact.

Having considered the negative impacts of commuting, it can also be a source of a positive utility. Commuting time is thus conceived as a gift rather than a burden. Depending on the transport mode used, scholars argue that there are opportunities for social and professional interaction and opportunities to experience the pleasures of the surrounding landscape or the social environment through gazing out of the window or watching other people, respectively. Longer commuting times also allow one to listen to music or the radio, to read, to relax, or to use the time more productively (e.g., work or study while traveling). With respect to the latter, the increased use of information and communication techniques (e.g., mobile phones, laptops, iPods, and Bluetooth technology) adds to the concept of new forms of copresence—that is, people are no longer restricted to connecting place to place but can now connect person to person. Sometimes commuting time is also experienced as “time out”—that is, time for being alone, time for doing nothing, or time to reinvest in maintaining social networks.

Apart from the economic, social, and health issues, there are also the spatial effects of the interaction of land use or urban form (characterized by density, diversity, design, and accessibility) on commuting behavior. For instance, a change in the location of the activities in which a person wants to participate and a change in the design characteristics of these locations will alter travel patterns. Therefore, integrating land use and transportation policy could alleviate transportation problems. Illustrative of this approach are the New Urbanism movement in the United States and the Compact City or Polycentricity Policy in Europe. These policies aim at modifying travel behavior (i.e., reducing average journey distances, trip frequencies, traffic volumes, energy consumption, and/or transport emissions) through land use planning. The basic idea is that high-density and mixed-use neighborhoods are believed to be associated with shorter trips and more non-motorized trips, hence less commuting.

Frank Witlox

See also Automobility; Exurbs; Mobility; Suburbs and Suburbanization; Transportation Geography; Urban Land Use; Urban Sprawl

Further Readings

COMPARATIVE ADVANTAGE

As the capitalist world economy unfolded, different regions increasingly came to specialize in the production of different types of goods and services. In Europe during (and often before) the Industrial Revolution, for example, Britain became a major producer of textiles, ships, and iron; France produced silks and wine; Spain, Portugal, and Greece generated citrus, wine, and olive oil; Germany, by the end of the 19th century, was a major exporter of steel, ships, and chemicals; Czechs were selling glass and linens; Scandinavia sold furs and timber; and Iceland exported cod to the growing middle classes. Within the United States, similarly, different places acquired advantages in some goods and not in others: The northeast was dominated by light industry, particularly textiles; the Manufacturing Belt became the center of heavy industry; Appalachia developed a large coal industry to feed the furnaces of the industrial core; the South grew crops such as cotton and tobacco; the Midwest became the agricultural products behemoth of the world; and the Pacific Northwest was incorporated into the national division of labor based on the expanding timber and lumber industry.

When regions, cities, or countries specialize in the production and export of some goods or services, they enjoy a comparative advantage. Although Adam Smith also used the idea to examine how firms specialized within a broader division of labor, this notion was spatialized by the famous 19th-century economist David Ricardo. Like all classical political economists, he assumed the labor theory of value (the value of goods reflects the amount of socially necessary labor time that goes into their production) and thus ignored demand. Ricardo concluded that nations will specialize in the production of a commodity that they can produce using the least labor compared with other nations.

Ricardo’s classic example of this process is demonstrated in Table 1, which illustrates the allocation of labor time in England and Portugal, two long-standing trading partners, before and after they specialized. In the first part, which depicts the labor hours per unit of wine or cloth that England and Portugal must each dedicate to the production of one unit of each good, it is evident that Portugal has an absolute advantage in both goods—that is, it can produce both of them with fewer labor hours than England can. If Portugal is more efficient, does it make sense for Portugal to trade? The answer is yes, implying that even the most efficient producer benefits from trade. Ricardo’s analysis examined what happens when each country allocates its resources to the good it can produce most efficiently compared with its trading partners—that is, when it acquires a comparative advantage. Thus, in the second part, England only produces cloth (two units at 100 hrs. [hours] each), and Portugal only produces wine (two units at 80 hrs. each). In the process of specializing—that is, of producing for a market that consists of both economies together rather than either of them alone, each country frees up some resources that would otherwise have been dedicated to the inefficient production of a good in which it did not have a comparative advantage. England saves 20 labor hours, Portugal saves 10, and the combined trading system thus saves 30, which can be reallocated toward investment (although the original model is static and says nothing about change over time).

The Ricardian model—the simplest of many, more complex notions of comparative advantage—has important implications for economic geography. First, it shows how powerfully trade and exchange shape local production systems. It demonstrates that trade allocates resources to the most efficient (i.e., profitable) ends. The only costs of free trade are those borne by inefficient producers—in this case, English wine makers and Portuguese textile producers. Second, Ricardian notions of comparative advantage reveal that specialization reduces the total costs of production; thus, trade improves efficiency even without reallocation of resources. Indeed, a good part of the long-term productivity growth of capitalism may be attributed to an increasingly specialized division of labor. Thus, specialization can be as important as technological change in stimulating productivity growth. For this reason, the vast majority of economists favor free trade as beneficial to all parties concerned. Third, this approach points out that large markets allow more specialization than do small ones. Adam Smith noted
the same thing when he famously stated that the division of labor is governed by the size of the market. In this case, when the market expanded from one country to two, it allowed firms to specialize and become more efficient in the process.

More generally, large countries with substantial internal markets facilitate more specialized firms and lower prices than do smaller countries.

Just as there is no specialization without trade, likewise there can be no trade without transportation. Goods must be moved across space from producer to consumer, and these transport costs must ultimately be borne by those who consume them. To the degree that transport costs affect the delivered price of commodities, they also influence consumers’ willingness to buy them and, thus, the competitiveness of the regions that export them. If transport costs are low, their impacts on the division of labor will be low. However, sometimes, particularly for heavy and bulky goods, transportation costs may increase the market prices of exports and imports prohibitively—that is, make the goods too expensive to ship profitably across regions. Throughout the history of capitalism, declines in transport costs have made it progressively easier for regions to realize their comparative advantage; thus, lower transport costs have contributed to lower production costs. For example, New Zealand became a major producer of lamb and mutton following the introduction of refrigerated shipping in the late 19th century. Similarly, the Pacific Northwest began to export vast quantities of wood and paper to the cities of the Midwest and the east coast following the completion of the transcontinental rail lines in the 1890s. In the post–World War II era, containerization likewise has facilitated the rise of various comparative advantages worldwide.

Ricardo’s two-country, two-product theory of comparative advantage can be expanded by allowing several production factors. The multi-factor approach to trade theory derives from the work of two Swedish economists, Eli Heckscher and Bertil Ohlin. The Heckscher-Ohlin theory holds that a country should specialize in producing those goods that demand the least from its scarce production factors. Unlike the original Ricardian model, it includes demand and allows for the production of more than one good. In this formulation, specialization of production will be incomplete—that is, countries may continue to produce some part of a good even if they do not enjoy total economic superiority in the costs of production. The Heckscher-Ohlin theory argues not only that trade results in gains but also that wage rates will tend to equalize. The reasoning behind this factor-price equalization, as it came to be called, is as follows: If a country specializes in a labor-intensive good, its abundance of labor diminishes, the marginal productivity of labor rises, and wages increase. Conversely, if another country specializes in capital-intensive goods, labor becomes less scarce, the marginal productivity of labor falls, and wages also fall.

The traditional theory of comparative advantage is simplistic and unrealistic. Ricardo never gave an adequate account of why regions specialize in some goods and not in others, instead offering a picture that is static with respect to time, overemphasizes labor and climate, ignores consumption as well as the role of economies of scale and agglomeration, says nothing about the nature of competition, and is silent concerning the

Table 1  Ricardo’s model of comparative advantage.
By revealing the gains created when England and Portugal traded, Ricardo’s model illustrates how specialization leads to more efficient resource use.

impacts of public policy. These shortcomings were addressed in the theory of competitive advantage.

Barney Warf

See also Competitive Advantage; Division of Labor; Economic Geography; Trade

Further Readings


**COMPETITIVE ADVANTAGE**

An improvement on the traditional theory of comparative advantage is the theory of competitive advantage. Unlike the Ricardian model, which was useful for understanding the simpler economies of the early period of the Industrial Revolution, this approach focuses on the social creation of innovation in a rapidly changing, knowledge-based economy. The key to competitiveness in this view is productivity growth: Over the long run, rising productivity creates wealth for everyone, if not equally. Productivity growth in turn reflects many factors, including the education and skills of the labor force, the available capital and technology, government policies and infrastructure, and the presence of scale economies.

Competitive advantage is dynamic and changes over time. The goal of national development strategies is to move into high value-added, high-profit, high-wage industries as rapidly as possible. To accomplish this goal, firms and countries should seek to sell high-quality goods at premium prices in differentiated markets. Quality is a key variable here; countries often acquire reputations for producing high- or low-quality goods, earning (or not earning) brand loyalty as a result. By moving into high value-added goods, nations seek to automate the low-wage, low-skill functions and retain the knowledge-intensive ones.

Although the global economy is increasingly seamless, competitive advantage is created in highly localized contexts—that is, within individual metropolitan areas. Globalization therefore does not eliminate the importance of a home base. Thus, countries that succeed internationally do so because a few regions within them move into “cutting-edge” products and processes. Within the United States, propulsive regions include Silicon Valley, Boston’s Route 128, and New York, with its premier position in finance and producer services; in Europe, they are Italy’s Emilia-Romagna, the continent’s largest high-technology region, as well as Germany’s Baden-Wurttemberg, Denmark’s Jutland peninsula, and the Cambridge region of the United Kingdom; in Japan, the government has actively constructed a series of technopolises toward this end (e.g., Toyota City).

The overall determinants of competitive advantage include the following: skilled labor, good educational systems, and technical training; agglomeration economies, including pools of expertise, webs of formal and informal interactions, trust, linkages, strategic alliances, trade associations, and integrated networks of suppliers and ancillary services; and a culture that rewards innovation—adaptation, experimentation, risk tolerance, and entrepreneurship—which includes heavy levels of corporate and public research and development and the continual upgrading of capital and skills. Corporations must engage in ongoing and organizational learning, anticipating changes in markets and demand; rigid corporate bureaucracies, like public ones, lead to complacency and short planning horizons. Competitive firms operate within competitive domestic markets; uncompetitive markets (i.e., private or public monopolies) exhibit little innovation. In the world economy today, increasingly sophisticated buyers spur a constant upgrading in the quality of output; adequate financing and venture capital; public policies that encourage productivity growth, including subsidized research, export promotion, and educational systems; and an up-to-date infrastructure (i.e., airports, telecommunications).

The theory of competitive advantage, largely attributable to Michael Porter, concludes that four attributes of a nation combine to increase or decrease its global competitive advantage and world trade:
1. Factor conditions
2. Demand conditions
3. Supporting industries
4. Firm strategy, structure, and competition

Factor conditions, or production factors, include human resources (the quantity of labor, skills, educational level, human capital, productivity, and cost of labor), physical resources (raw materials and their costs, location, access, and transport costs), capital resources (funds to finance the industry and trade, including the amount of capital available; the savings rate; the health of money markets and banking in the host country; government policies that affect interest rates, savings rates, and the money supply; levels of indebtedness; trade deficits; and public and international debt), knowledge-based resources (research, development, the scientific and technical community in the country, its achievements and levels of understanding, and the likelihood for future technological support and innovation), and infrastructure (all public goods and services necessary to facilitate the production of goods and services that provide a country with a competitive advantage). Also included are transportation systems; communications and information systems; housing, cultural, and social institutions; and education, welfare, retirement, pensions, and national policies on health care and child care. These five factors are identified in current international and economic circles as the keys to the competitive advantage of a nation in the foreseeable future.

Demand conditions are the market conditions in a country that aid the production processes in achieving better products, cheaper products, scale economies, and higher standards in terms of quality, service, and durability. Demand conditions cause firms to become innovative and, therefore, produce products that will sell not only in the domestic market but also in the world market.

To be competitive internationally, firms require easy and continuous access to networks of other firms specialized in different tasks. For example, large financial institutions require law firms, marketers, computer services, accountants, advertisers, and management consultants, subcontracting tasks that require heavy investments in human capital. Access to these industries, which generally provide specialized expertise, is often obtained through face-to-face contact.

Firm strategy, structure, and competition relate to the conditions under which firms originate, grow, and mature. For example, because stockholders demand U.S. companies to show short-term profits, U.S. corporate performance may be less successful in the long run than it would be if it were judged over a longer period of time, as Japanese and German corporate performance is.

State support of corporate strategy and performance is important. For example, a country can regulate taxes and incentives so that investment by a firm is high or low. In addition, competition within a country can impose demands on company performance; new business formations often pressure existing firms to improve products and lower prices and thus increase competitiveness.

Barney Warf

See also Clusters; Comparative Advantage; Economic Geography; Innovation, Geography of; Trade

Further Readings
distributed artificial intelligence, in which entire societies of intelligent agents interact with one another to simulate the interaction of individuals in large populations. The geographer Michael Batty explains that a complex system is based on simple rules or interactions that give rise to unanticipated spatial outcomes. John Holland, a pioneer of complexity and nonlinear science, notes that a complex adaptive system (CAS) is a dynamic network of many agents (which may represent cells, species, individuals, firms, or nations) acting in parallel, constantly acting and reacting to what the other agents are doing. The control of a CAS tends to be highly dispersed and decentralized. If there is to be any coherent behavior in the system, it arises from competition and cooperation among the agents themselves, as opposed to being orchestrated by a higher-level authority or structure. The overall behavior of the system is the result of a large number of decisions made every moment by many individual agents. Thus, order in a complex system is understood as an emergent property of individual interactions at lower levels of interaction.

These fairly general definitions of complexity science bear an uncanny resemblance to those offered by general systems theory (GST), introduced to geography in the 1960s as part of the quantitative revolution to provide a framework in which to understand and theorize various disparate strands of work. GST was based on the work of Ludwig von Bertalanffy, who claimed that GST allowed for the transfer of principles discovered in one context to other situations. Although dismissed by Nicholas Chisholm in 1967 as an “irrelevant distraction,” the themes of GST continue to haunt geography, coming to the forefront once again in the dialogue between Jonathon Raper, David Livingston, and Doreen Massey in the late 1990s, as they discussed what Massey (1999) called the “commonalities between physical geography and human geography in emerging ways of conceptualizing space, time, and space-time” (p. 261). David O’Sullivan joined this conversation in 2004, when he brought it into the complexity arena, noting that “research in the complexity modus operandi can be useful, which brings us back to the potential for intra-disciplinary discussion provided by the complexity venture” (p. 283). The distinguishing element between general systems theory and complexity theory is the reconceptualized understanding of equilibria in complexity theory. Chaos theory, and later complexity theory, allows for small shifts in system parameters to produce positive feedback loops, pushing the system to new levels of equilibria as opposed to dampening it through negative feedback.

The first application of complexity theory to geography may have been the segregation experiments done by the Nobel Laureate Thomas Schelling in 1971, who found that he could show how even the most racially and ethnically tolerant urban residents produce a highly segregated landscape in their individual search for a home. Batty may now lead the field in his application of complexity theory to urban systems, ultimately noting that “the city [is] the example par excellence of organized complexity” (p. 4). Steven Manson seeks to define complexity and its role in geography, detailing three categories of research where the term is used most frequently: algorithmic complexity, deterministic complexity, and aggregate complexity. While the first and second remain the purview of computer science, the third, aggregate complexity, has potential ramifications in geography. Defined as “the study of phenomena characterized by interactions among distinct components” (p. 409), this understanding or interpretation of complexity is of particular interest to geographers as it implies that the local spatial configuration of interactions affects outcomes at the whole-system level. Labeled emergence in the computer science and more popular literature, this phenomenon, also called self-organization, is a central feature of complex systems.

While complexity science remains fairly poorly defined, the tools that it has spawned are potentially powerful in geographic research. Most notably, multiagent systems, the newest and most complicated of a cadre of tools available, are composed of a set of agents and sets of cells or layers that define different landscapes or environments. Agents do not necessarily have fixed locations, and they interact with one another as well as with the environment in which they reside. These agent-based models allow social scientists to engage in thought experiments related to real-world phenomena. Multiagent systems are inherently geographic, due to their fundamental
understanding and modeling of the interaction between agents and the environment.

Unlike research in general systems theory, research taking place under the rubric of complexity theory may produce substantive achievements. Batty has published several articles modeling the temporally dynamic features of spatial behavior—work that falls within the purview of complexity as defined here. While a significant portion of that work discusses the tools developed to simulate social-spatial behavior, Batty also has contributed to the theorization of cities as fractals—once again searching for universals in the most particular. Dawn Parker, an agricultural economist now making tremendous contributions to the geography literature, has written extensively of multiagent systems in geographic research, especially in understanding land use and land cover change. Although Parker does not identify herself as a “complexity theorist,” the methods she has worked to develop and introduce to geography fall under the umbrella of complexity, most notably the multiagent systems. Parker’s models integrate decision-making agents (e.g., households and land parcel managers) working within social networks, markets, and political institutions, to influence the environment in the form of soil fertility, erosion, and the spread of invasive species. As opposed to previous techniques used to model land use and land cover change, Parker’s models are simulation, rather than equilibrium, based, once again putting her work clearly in the rubric of complexity, recognizing the nonlinear nature of these intensely interactive systems prone to positive feedback rather than equilibrium.

Christine Drennon

See also Agent-Based Models; Cellular Automata; Complex Systems Models

Further Readings


Rise of the Concept of Complex Systems: From Analysis to Synthesis

The triumphant progress of physics aided by mathematics since Galileo Galilei is associated with a major focus on linearity. Relying on the
assumption of linearity that a whole is the sum of its parts, our combined mathematical tools and physical knowledge have made a tremendous scientific contribution. Such an approach, for hundreds of years, produced a deep-rooted reductionist tradition characterized by an analysis of repetitively fractioned units (e.g., cell > molecular > atom > nucleus) rather than a synthesis of those components in the framework of holism.

The science of complexity originated from the recognition of complex systems whose behavior cannot be understood based only on the reductionist approach. For example, the study of individual cancer cells in detail provides a limited insight into how they interact and spread. Intensive knowledge on local atmospheric conditions across the planet, without integrating the discrete localities into a whole, would not lead to an understanding of the mechanism of global warming. Urbanization results from combination of a demographic process by which the size of a population and its concentration increase and economic processes in which its economic base grows. In addition to these examples of complexity, the origin of life, biogeochemical cycles, forest fragmentation, stock market gyrations, traffic jams, and rising riots from a silent crowd remain mysteries.

One of the underlying processes in the behavior of these complex systems is nonlinearity, which makes a whole larger or smaller than the sum of its parts. A system is considered complex not because it consists of multiple variables but because their interactions and associated processes are nonlinear. In CAS, a considerable number of components interact, continuously changing their state from moment to moment, which is a system evolving through self-organization. The system cannot be investigated by isolating any one part because the dynamic interactions lead to positive or negative feedback loops. In other words, a certain microlevel property can extend its control to a characteristic phase of the synoptic level at which the system is perceived as a whole, and the synoptic level pattern in turn controls the microlevel behaviors. At the microlevel, nonlinear dynamics in some cases leads to chaotic behavior by which variations in the initial conditions are magnified over time, which thereby hampers any precise predictability. Through the feedback process, however, the synoptic level of self-organization can balance this microlevel chaotic behavior, producing a fractal pattern that is self-similar across levels.

A CAS is also characterized by emergent properties. They generate spontaneously without any defined rule during the process of self-organization in CAS and cannot be inferred if one simply aggregates the individual components. They maintain themselves as an unexpected, integrated whole for some time period. Emergent properties are therefore perceived as novel, coherent structures. In his model of spatially divergent self-organization on the Earth’s surface as nonlinear dynamical systems, Jonathan Phillips suggests that self-organization involves the progressive differentiation of landscapes (i.e., emergence) from more homogeneous to more heterogeneous patterns. Here, a landscape is composed of sets of interacting components, and such mutual interactions can explain its self-organization without taking into account the direct effects of external forcing. He postulates that entropy serves as an appropriate measure of the organization in the following model:

$$\delta(t) = C e^{\lambda t},$$

where $\delta(t)$ is a perturbation at time $t$, $C$ is a constant or initial variation at $t = 0$, and the arrows indicate vector quantities. $\lambda$ is a Lyapunov exponent. The number of Lyapunov exponents is the same as the number of components in a system. In a mathematical sense, deterministic chaos is defined when there is at least one positive Lyapunov exponent, thus $\lambda > 0$ implies an exponential increase of variation and the unstable, divergent behavior of a landscape. The sum of positive Lyapunov exponents is known to represent the Kolmogorov (K-) entropy, or Kolmogorov complexity. It is therefore summarized that there is a local connection among nonlinear chaotic dynamics, the Lyapunov exponent, and K-entropy. For example, if the initial spatial variation of nutrients in a landscape is magnified over time (i.e., there is nonlinear chaoticity), vegetation cover would differentiate into diverse spatial units (i.e., result in landscape divergence). Such divergence involves positive Lyapunov exponents, thus positive K-entropy.
COMPLEX SYSTEMS MODELS

Complex Systems Models in Geography

Because the key to the science of complexity lies in the synthesis of the knowledge and techniques of various research fields, the study of CAS necessitates interdisciplinary cooperation. Geography in itself is multidisciplinary, so many geographical models have the potential to contribute to the progress of complexity theory. This multidisciplinary nature also implies that the complex systems models used by geographers are in most cases developed and improved in cooperation with adjacent fields.

Conceptual Models

Conceptual models are concise and simplified visualizations that illustrate how complex interactions among system components generate patterns (e.g., Figure 1). Compared with mathematical models, conceptual models tend to be less appreciated explicitly in the discussion of complex systems models.

Conceptual models are useful because they do the following:

1. They enhance an understanding of the mechanisms underlying systems operation, thus often becoming a basis for predicting complex interactions across space and time, under unprecedented scenarios.

2. They generate hypotheses that can be tested by experimental studies or mathematical modeling. These two approaches in turn help improve the conceptual models themselves.

3. They facilitate interactions among academics, policymakers, and the public because conceptual models express technical, complex details in a nontechnical, visually stimulating way.

4. They encourage interdisciplinary cooperation. Conceptual models involve a truly intensive process of synthesizing a variety of knowledge, which is hardly achieved without integrating perspectives from various disciplines.

Mathematical Models

Most mathematical models of complex systems are based on a common premise that simple underlying interactions are able to generate complex spatial and temporal patterns. This premise also implies that any complexity observed in natural and human systems does not necessarily involve complex causes. Focusing on complexity in time and space, three major categories of complex systems models are discussed below.

Discrete Temporal Models

Until the 1980s, when space was not often explicitly involved in mathematical and conceptual models of complex systems, discrete, non-overlapping temporal dynamics with the deterministic chaotic behavior of a system had been modeled predominantly using difference equations:

\[ N_{t+1} = F(N_t) = \mu N_t (1 - N_t) \]

The logistic equation above represents population density \( N \) and the dynamics between generations \( t \) and \( t + 1 \), with \( \mu \) being the population growth rate. As the growth rate \( \mu \) increases, the population density shows a variety of dynamical long-term bifurcations (Figure 2). Chaos theory is typically associated with the “butterfly effect,” in which even slight changes in the initial system conditions stimulate nonlinear interactions among multiple elements to produce dramatically different final states.

Based on these random-like, aperiodic bifurcations, some considered chaotic behaviors as maladaptive or catastrophic since the violent unstable fluctuations may lead some population values to approach an extinction threshold as shown in Figure 2. However, such an argument does not take into account the spatial component in complex systems.

When space is incorporated into temporal chaotic models, for example, by introducing the concept of dispersal, the argument against chaos can be resolved. The stabilizing role of space explains that dispersal between local populations facilitates the persistence of metapopulations at the global spatial scale, while there may be extinctions at a local spatial scale. This global
Figure 1  A conceptual model of coastal dune ecology synthesizing spatial gradients of vegetation, soil, and geomorphology. The model illustrates how spatial patterns of vegetation are related with soil conditions controlled by landform and geomorphic processes. Such a conceptualization can be of particular relevance in coastal ecosystems, where the dynamic interactions among atmospheric, marine, and terrestrial systems hamper empirical modeling.

stability is termed chaotic stability. The next two sections discuss in greater depth complex systems models that couple space and time.

**Continuous Space-and-Time Models**

Continuous spatiotemporal models are equivalent to the so-called reaction-diffusion models in physics, simulating a system in which local quantities interact according to basic physical laws. One classic example is found in chemical reactions, in which a number of molecules (i.e., local quantities) diffuse across space and, in case of close contact, react with one another. In the field of geography, various spatially distributed processes, such as urban growth, land use change, and population dynamics that are influenced by the spatial structure of the environment, transport infrastructures, and movement patterns, can be investigated using the reaction-diffusion models.

In metapopulation epidemic models, for instance, diffusion is explained by the migration of people between cities, each of whom is classified into different states such as infected, susceptible, or immune. Here, reaction represents the probability of individuals in the same city being physically close enough to one another to change their state of disease.

In its simplest form, the model of diffusion-reaction is expressed by the following equation of population dynamics:

\[
\text{Change in population} = \text{Growth of population} - \text{Diffusion of population}.
\]

This elementary equation explains that a city grows through the growth of its population and spreads by the diffusion of the population in the suburbs. Here, some suggest that the equation should be generalized by taking into account the
advection processes by which a demographic excess is induced through the immigration of people from the rural areas or other cities. Such a suggestion holds that a city’s growth and expansion are the result not only of the auto-growth of its population but also of the external input, as illustrated by the following revised equation.

\[
\text{Change in population} = \text{Growth of population} - \text{Diffusion of population} + \text{Advection of population}.
\]

One further generalization can be attempted for this model by coupling two or more equations or adding interaction terms. The population of the model, for instance, may consist of various ethnic groups or economic classes dynamically interacting with one another, as can be observed in real society. If one is interested in any two group types, the equation can be extended as follows:

\[
\begin{align*}
\text{Change in Population A} &= \text{Growth of Population A} - \text{Diffusion of Population A} + \text{Interaction of Populations A and B} \\
\text{Change in Population B} &= \text{Growth of Population B} - \text{Diffusion of Population B} + \text{Interaction of Populations B and A}.
\end{align*}
\]

In a biogeographical sense, Populations A and B can be two different plant species competing or facilitating each other.

To model geographical phenomena adopting the theory of diffusion-reaction, the theory should be enriched by the consideration of two aspects: spatial heterogeneity and human agents. Recently, a body of literature began to emphasize network heterogeneities across space, doubting the previous assumptions of spatial homogeneity, isotropy, and randomness that the classic diffusion-reaction models are implicitly based on. Geographers are concerned with human and natural processes occurring on the Earth’s surface, which is always structured and heterogeneous. On a beach, sand grains are transported by onshore wind (i.e., diffusion), interacting with one another during their saltation (i.e., reaction). Here, the underlying topography is the fundamental control of the path of such transport (i.e., spatial heterogeneity). In a city, population changes by the emigration or immigration of individuals predominantly along road networks (i.e., diffusion and advection). Underlying preliminary structures such as topography and land use constrain the spatial pattern of the networks, thereby making them spatially anisotropic and heterogeneous.

Second, the classic diffusion-reaction model itself cannot be fully compatible with geographic explorations because humans are not like physical and chemical entities such as molecules and atoms. Humans’ behavior reflects their personal and societal needs, which involve emotional, cultural, and economic aspects. A challenge is then how to incorporate such original human actions into the physical diffusion-reaction models. The most important point for the human agent is that humans, as social species, interact with each other, thereby comparing their cultural and economic status with their neighbors’ situation. Modeling studies have shown that the segregation of homogeneous quarters emerges from a random initial distribution of people who tend to continuously diffuse toward a place where they can share similar cultural values.

**Discrete Space-and-Time Models**

Discrete spatiotemporal models are classified into coupled map lattices (CMLs) and cellular automata (CA) models. The main purpose of a CML is to model the spatiotemporal chaos created by local nonlinear dynamics and spatial diffusion in a system with discrete space, time, and continuous state variables. In a CML, space is represented by a two-dimensional lattice in which the state dynamics of each point are defined as real numbers or real vectors. The definition here is achieved by coupling each site to its nearest neighbors, assuming that local interactions exist between the neighbors and the point in the middle. The simplest coupling is called the Von Neumann neighborhood, in which the state of a point at time \( t + 1 \) is an interactive function of the states of its four cardinal neighbors (i.e., north, south, east, and west) at time \( t \). In square-lattice neighborhoods, this simple coupling can be extended to the Moore neighborhood, which considers all eight surrounding points. In many ecological systems and population dynamics models, where clustering of points or patches is important, a global
coupling is often used when interactions occur beyond the spatial extent of direct neighbors (i.e., the Moore neighborhood). In a formal way, a CML is expressed by the following equation:

\[ x_{n+1}(i) = (1 - \varepsilon)f(x_n(i)) + \frac{\varepsilon}{2} \left[ f(x_n(i+1)) + f(x_n(i-1)) \right], \]

where \( i \) is the spatial location of a point with a periodic boundary condition. In this basic model, the dynamics of a CML is determined by two competing components, such as individual non-linear reaction (i.e., \( f \)) and spatial interactions (i.e., coupling) with intensity \( \varepsilon \).

Since their introduction to the field of physics and biology by Kaneko, Kapral, and Kuznetsov in the early 1980s, CMLs have been used by some geographers and others from related fields. One of the models’ benefits is that CMLs are ideal for computer simulations. Traditionally, studies on the dynamics of spatially extended systems had been conducted under various discretization schemes required by partial differential equations. However, a direct numerical simulation without such schemes became possible using CMLs because CMLs are discrete in space. This advantage has been crucial for facilitating a variety of numerical simulations, thereby generating great insights into the spatiotemporal patterns of different regimes of dynamics.

Another group of discrete spatiotemporal models are the CA. Different from CMLs, where states are continuously defined as real numbers, CA are discrete in state. In a biogeographic vegetation dynamics model, for instance, the state of a cell can be defined as bare ground (0), pioneer stage (1), midsuccessional (2), or late successional (3).

Daehyun Kim

---

See also Agent-Based Models; Cellular Automata; Complexity Theory; Models and Modeling

---

Further Readings


The CLAG is a nonprofit organization governed by a board of directors elected by the general membership, which has varied in number between 200 and 400. There are 13 members on the board, each serving for 3 yrs. (years). A rotation policy brings new members to the board each year and retires members who have completed their terms. Officers of the board are the chair, vice chair, and executive director, who may be assisted in their duties by appointed staff. Committees of the board are appointed by the chair to undertake organizational business; they include an executive committee as well as honors, membership, program, publications, teaching, and research committees. The executive committee is responsible for the integration and facilitation of the activities and plans of the organization.

CLAG meetings are scheduled at 1- to 2-yr. intervals. These meetings, which have been held in Latin America, Canada, and Spain, as well as the United States, consist of volunteered papers on a general conference theme, with a keynote address by a noted Latin Americanist. The first such meeting in 1970 was funded by the National Science Foundation, the Social Science Research Council, and the American Council of Learned Societies.

The CLAG annual business meeting, open to all members, is held in April each year. One meeting at the beginning of each decade is devoted to an inventory of geographical research in Latin America in the preceding 10 yrs. and a discussion of prospects for the future.

During the 1980s, CLAG made great progress in becoming a truly international organization; indeed, only two of CLAG’s eight meetings since 1981 have been located in the United States; Latin American sites have included the Dominican Republic, Mexico, Colombia, Honduras, Peru, Costa Rica, Spain, and Guatemala.

Printed and electronic publications include the annual, peer-reviewed *Yearbooks* (1970–2001), now replaced by the two-issue-a-year *Journal of Latin American Geography*, special publications, occasional publications, instructional media, and the *CLAG Newsletter*. CLAG also maintains its own listserv.

David J. Robinson

*See also* Association of American Geographers

---

### Further Readings

Conference of Latin Americanist Geographers: [http://sites.maxwell.syr.edu/clag/clag.htm](http://sites.maxwell.syr.edu/clag/clag.htm)

---

**CONFLATION**

The word *conflation* refers to the process (compilation, reconciliation) or the resulting product of combining two or more versions of a complex information entity, such as a manuscript, map, or database, to produce a new improved version of the same entity. A map conflation is the process or the output map resulting from merging information from two or more maps of the same region. Spatial data conflation is the process of, or the output data set or database resulting from, merging two or more spatial data sets or databases for the same area or region. Although the terms *conflation* and *data fusion* are sometimes used interchangeably, data fusion is technically the more general concept of merging any two data sets, whereas conflation merges data sets that are presumed to each contain large subsets of common features. Map conflation is considered successful only after many or most features on one map have been matched to counterpart features on the other.

The first prototype semiautomated, computerized map conflation system was built in 1985 for a study of the feasibility of merging urban map data (Digital Line Graph [DLG] files) from the U.S. Geological Survey (USGS) with urban map data (Geographic Base File/Dual Independent Map Encoding [GBF/DIME] files) from the U.S. Bureau of the Census. At that time, long before the current widespread proliferation of spatial data sets, only government agencies had access to multiple map data sets of large coverages; hence, only government agencies had the opportunity, the resources, and the incentive to try to develop computer programs to conflate such data sets.

The USGS/Census semiautomated map conflation system employed a *match-and-merge conflation strategy*: Find duplicate features in the two
map versions, then display all the unique features (duplicated or otherwise) exactly once to create a conflated map of the two versions. More specifically, the steps are as follows:

1. Using a variety of map-feature similarity measures, find candidate matching pairs of features (every pair will have one element from each of the two map versions).

2. Apply a topology-preserving geometric transformation (commonly referred to as rubber sheeting) to all the points of one of the map versions to bring candidate-matched features into exact alignment with each other.

3. Simultaneously display the two realigned and newly registered map versions for visual assessment of the effect of collocating the candidate-match pairs and displaying each pair as a single feature.

Since 1985, many private companies, academic institutions, and government agencies have continued developing spatial data conflation software. Although the matching criteria and alignment transformations vary slightly from system to system, all systems use the two-part feature-matching/feature-alignment model to generate graphic displays of the results of the conflation processing.

Map-to-map conflation has been expanded to include map-to-image conflation and image-to-image conflation. Working with images requires an extra step of image feature recognition, but once image features have been identified, those features can be matched and merged (i.e., rubber sheeted) to align with map features or other image features using previously developed matching and rubber-sheeting tools.

---

**Further Readings**


---

**CONSERVATION**

The term *conservation* usually refers to environmental conservation or preservation of natural resources. In the United States, the term came into use in the late 19th century, when valuable natural resources such as timber, farmland, and pasture were threatened by rapid economic development. Concerns extended to include the Great Plains as frontier settlement moved westward at an increasingly rapid rate with the construction of continental railroads, which laid land open for further development. The term suggests preservation and protection of natural resources to prevent exploitation and destruction. Conservation also includes the recognition of limits to growth. Presently, the term conservation embraces the ecosystems and resource base of the Earth by protecting its capacity for self-renewal and sustainability.

---

**The Conservation Movement in the United States**

The conservation movement has a long history in the United States. In the 1850s, Henry David Thoreau—the author of *Walden*—expressed his concern about the loss of nature and saw in the “wild” the hope for the preservation of the world. He called attention to the sacred laws of nature, which take precedence over human actions that disturb the ecological balance. In 1864, George Perkins Marsh formulated the physical geography of the Western states as modified by human action and drew attention to the fragility of the natural ecosystem. These early efforts of consciousness-raising led to a concerted effort to protect nature and establish national parks in the Western United States. It was President Theodore Roosevelt...
Roosevelt, with the aid of his chief forester, Gifford Pinchot, who established the federal role in these efforts. Roosevelt began withdrawing large areas of western public land from further development, and in 1916, the National Park Service was created to preserve valuable nature for eternity. The drought conditions, known as the Dust Bowl of the 1930s, led to the foundation of the Civilian Conservation Corps in 1933 and the Soil Conservation Service in 1935 during the presidency of Franklin D. Roosevelt. Conservation efforts in the United States also include concerns about water resources, energy use, and water and energy conservation.

In the 1960s and 1970s, the conservation movement became more actively involved in the political debate about pollution and the destruction of nature in favor of economic profit. Social justice issues transformed the conservation movement into the environmental movement and engaged the political establishment in protecting resources and the environment. The environmental movement produced profound changes in the political climate concerning the environment, culminating in the establishment of the U.S. Environmental Protection Agency in 1970.

The Whole Earth Ecosystem

In the late 1960s and early 1970s, concerns about the health of the global ecosystem and resource base found resonance with Rachel Carson’s message about the detrimental widespread use of pesticides in agriculture. In the same period, Paul Ehrlich and Barry Commoner raised issues about population, economic growth, and technology as causal factors in the limits-to-growth debate. Ehrlich presented a neo-Malthusian scenario of imminent population explosion and ensuing disaster and argued that society must find ways to limit population growth. Commoner presented an eco-socialist perspective to the limits-to-growth debate, arguing that capitalism and technology were chiefly responsible for the environmental degradation of Earth’s ecosystem.

Conservation and sustainable development are closely related, and the integral relationship between the health of Earth’s ecosystem and human well-being is widely recognized. Global wildlife conservation and biodiversity are high on the agenda of the global conservation movement and find supporters in many countries, but equally important are concerns about fossil fuel use and greenhouse gas emissions. Protection of the global environment and conservation of Earth’s resources represent an enormously complex challenge at a time of rapid globalization. The advancement in science and technology, communication, and transportation that has united the world physically has also greatly accelerated the threat of reduced biological diversity and climate change. Materialism, driven by the dogmas of consumerism, capitalism, and individualism, has in many instances been carried to excess and is now affecting livelihoods in many parts of the world. Extending social and economic justice to the entire global community and ensuring quality of life for subsequent generations are important principles underlying the global environmental conservation movement.

Ecological and Carbon Footprints

Concern about sustainable development and the threat of climate change has focused attention on humanity’s ecological and carbon footprints. The political challenge of responding to the ecological crisis and climate change is daunting. The ecological footprint measures how much land and water area a human population requires to produce the resources it consumes and to absorb its waste products. Today, the ecological footprint is almost 25% larger than what the Earth can regenerate in 1 year. By measuring the ecological footprint of a population (i.e., of an individual, a city, a nation, or all humanity) one can assess overshoot, which should help in managing the ecological assets more carefully.

The carbon footprint is a subset of the ecological footprint and expresses the impact human activities have on the environment in terms of the amount of greenhouse gases produced, measured in units of carbon dioxide. It is meant to be useful for individuals, nations, and organizations to conceptualize their personal (or organizational) contribution to global warming. It is estimated that developed countries’ average per capita carbon footprint is 10 times that of the average developing countries and that to attain “safe” levels of carbon dioxide concentrations...
in the atmosphere, carbon emissions need to be reduced by 60% worldwide. To reduce the carbon footprint, alternative energy sources such as solar and wind energy need to be rapidly developed, and energy conservation has to be widely practiced.

Yda Schreuder

See also Biodiversity; Biosphere Reserves; Commons, Tragedy of the; Community-Based Conservation; Conservation Zoning; Ecological Footprint; Ecosystems; Environmental Management; Governmentality and Conservation; Landscape and Wildlife Conservation; Neo-Malthusianism; Patches and Corridors in Wildlife Conservation; Population and Land Degradation; Renewable Resources; Resource Economics; Resource Geography; Resource Management, Decision Models in; Soil Conservation; Spatial Strategies of Conservation; Sustainable Development

Further Readings


Conservation zoning is a spatial planning strategy used to manage urban and rural landscapes and sometimes marine areas. The term can refer to any land use or marine planning process whose primary goal is purported to be the conservation of social and/or natural resources. It is a territorial mechanism that delineates bounded zones for which particular rules apply—in this case, rules meant to conserve particular environmental values. At its most basic, conservation zoning sets out acceptable density levels and residential footprints to ensure the persistence of green space. In urban landscapes, both social and environmental characteristics are targets of conservation zoning, whereas in rural areas, conservation zoning is used primarily to protect natural resources and ecosystem services. Geography and geographers have played a role in determining the types of conservation zones appropriate to a planning exercise and, through the use of spatial technologies and multicriteria decision support, have been integral to the establishment and mapping of conservation zones.

While residential, industrial, and commercial zoning are common practices in urban land use planning, conservation zoning is a relatively new addition to the urban-planning toolbox meant to promote the building of more sustainable communities. Urban conservation zones range from protected heritage architecture to riparian buffer zones that border fish-bearing streams and their floodplains, as well as urban green spaces. Commonly, conservation zoning limits particular kinds of development, such as land clearing and construction, that may threaten the targeted social or environmental characteristics.

In rural and peri-urban areas, conservation zoning is a way of promoting land-based conservation values on private property. Conservation easements and land trusts are ways in which property owners can receive municipal tax incentives in exchange for preserving their land to ensure the provision of habitat and/or ecosystem services. Greenbelts or greenways are examples of conservation zones found most often bordering urban areas. Both restrict development to preserve forest, agriculture, or otherwise environmentally sensitive land. Parks and protected areas can also be thought
of as exercises in conservation zoning where core zones are delimited from multiple-use, recreational, or buffer zones.

In the vast oceans, especially the Pacific, conservation zoning is a key means of ensuring and promoting networks of marine-protected areas (e.g., Pahānâuaumokuākea, Marianas Trench, Rose Atoll, and Pacific Remote Islands Marine National Monuments of the United States).

The process of conservation zoning is political and requires the spatially explicit analysis of numerous data sets—everything from soil, land cover, slope, and connectivity to property values, public perception, recreational use and potential, and demographic characteristics. Geographic tools, for example, remote sensing and geographic information systems, are ideal for such analysis. Their use informs decision-making processes and helps identify lands suitable for conservation zoning.

Robin J. Roth

See also Biosphere Reserves; Conservation; Environmental Certification; Governmentality and Conservation; Greenbelts; Indigenous and Community Conserved Areas; Land Use Planning; Resource Management, Decision Models in; Spatial Strategies of Conservation; Zoning

Further Readings


Consumption has meant different things at different moments in time and has taken widely varying forms. The term comes from the Latin consumere, “to use up” (as in tuberculosis), implying that consumption means to make full use of and to destroy simultaneously. In societies lacking a social surplus, consumption equals production. Economically, consumption is how consumers satisfy their needs. In social terms, consumption lies at the intersections of production, culture, everyday life, and psychology. In contrast to production, which has been studied in exhaustive detail in geography, consumption has long been ignored or regarded as unproblematic. The reasons for this silence are not clear but may reflect, among other things, Marxism’s emphasis on production and labor as the central acts of social life and, conversely, neoclassical economics’ sterile and asocial view of consumption. Consumption and production cannot be neatly separated and are closely intertwined: Most people work in order to consume, and everyone consumes in order to live.

Historically, the growth of mass production in the 19th century was accompanied by rising disposable incomes, mass consumption advertising, the ideology of consumerism, and, in the 1930s, Keynesian demand management. Consumption fuelled by debt was decried by moralists, often on religious grounds, as a violation of Protestant ethics, which focused on delayed gratification. In the late 20th century, changes in the world economy, including deindustrialization and the explosive growth of producer services, induced concomitant changes in consumption, including increasingly specialized niche markets and sophisticated consumers. By any measure, consumption is enormously important, as an economic act (constituting the bulk of the gross national product [GNP] of most countries), environmentally (e.g., energy use or the act of turning products into trash), and in terms of the lifestyles and self-image of much of the population. The geography of consumption is critical to understanding related issues such as travel and transportation, tourism, standards of living, and uneven development.

Theoretical Perspectives on Consumption

The historically dominant view of consumption came from neoclassical economics, which analytically privileges demand. In this perspective, individual consumers, personified by the
desolate, self-centered, asocial character *Homo economicus*, maximize their utility or happiness by allocating incomes among different goods. This subject has been examined extensively, including topics such as the impacts of changing incomes and prices, consumer surplus, elasticities of supply and demand, and the impacts of imperfect information. Consumption forms the vast majority of gross domestic product, often 80% or more. Inevitably, the conclusion of such views is that markets are optimally efficient (and hence morally optimal as well). While the neoclassical view is internally consistent within its own terms of reference, it is ultimately sterile and ahistorical, failing to do justice to the rich semiotics and social dimensions of consumption. In part, this failure arises because neoclassical economics does not represent consumption as a social act—that is, embedded within broader relations of class, gender, ethnicity, and power. For example, it offers no account of the origins of utility curves or why they assume their particular form. Social categories, if they arise at all, are defined largely by their relations to consumption: class in conventional Weberian social analysis, for example, refers to income and socioeconomic status.

A second interpretation of consumption comes from Marxism, which argues that social science must penetrate the veneer of outer appearances to reveal the social relations that lie beneath them. In this vein, Marx argued that commodities are not simply *things* but embodiments of social relations. To view commodities separately from their social origins is to commit the error of commodity fetishism. The means by which market relations render relations among producers opaque is functional for capitalism, hiding the source of wealth, that is, labor. Rather, Marxism drew on classical economics to differentiate the use value of commodities—the qualitative, subjective dimensions—from their exchange value, the quantitative price they command on the market (e.g., the use value of an apple is its taste and the relief from hunger it offers; its exchange value is how much it sells for on the market). Critically, for Marxists, labor too is a commodity whose use value to employers is less than its exchange value in wages. Class is thus defined by relations to production, not consumption. Marxism suggests that the extraction of surplus value by employers inevitably leads to underconsumption by the working class and the tendency toward crisis. A third perspective on consumption focuses on the semiotic dimensions. Rather than a simple act of utility maximization, as represented by neoclassical economics, this body of work points to shopping and consumption as social and spatial practices that emanate from and in turn reinforce existing structures of power, culture, and ideology. Early sociologists, such as Thorsten Veblen, focused on the lifestyles of the robber barons and their propensity toward “conspicuous consumption” as a sign of wealth and status. In highly individualized societies such as the United States, personal status is typically achieved through the consumption of commodities. Indeed, self-identity and even self-esteem are frequently linked to owning the “right” brands of goods. What one may call the sensuous nature of consumption, thus, includes the complex social and psychological motivations that underpin the urge to buy, including consumers’ ego, sense of self, status definition, and alleged individuality that comes from the purchase of mass-produced commodities.

The highly influential early-20th-century cultural theorist, Walter Benjamin, extended historical materialism to include the bourgeois infatuation with the commodity. He sought to uncover the ways in which the commodity penetrated into the consciousness of buyers, charting the growth of bourgeois consciousness in the emerging malls and stores of early-20th-century Europe. Working in Paris and Berlin in the 1920s, Benjamin’s Arcades Project examined the linkages between the urban environment, experience, history, and memory, portraying cities as labyrinths in which individual subjectivity was swept aside by modernity and its impersonal relations, bureaucracies, and markets. Perception, Benjamin maintained, was itself historically specific. Commodities, in this reading, were far more than embodiments of labor power; they represented visual aesthetics with a significance above and beyond the narrow realm of the economic. In this light, Benjamin revealed that commodities are as much distillations of signs as they are embodiments of use and exchange values. Benjamin’s Arcades Project thus captured the commodified nature of modernity, the deep linkages between seeing and knowing, on the one
CONSUMPTION, GEOGRAPHIES OF

This step effectively opened up the analysis of consumption as a social process, noting its mounting autonomy from production and the pervasive role of symbols in the construction and manipulation of consumer consciousness. This line of thought reemerged in postmodern analyses of consumption, particularly the astute critique of contemporary capitalism offered by the political economy of signs. For Baudrillard, the mass media have made the sign more important than its referents, creating a world of the simulacra, in which we can no longer distinguish between simulations and reality, between the true and the false. In the context of post-Fordism, postmodern consumption centers as much on the symbolic value of commodities as their use value. Thus, pseudo-Irish bars are more Irish than Ireland. Baudrillard’s dissection of Disney World reveals it to be just such a simulacrum, a giant shopping mall, and for Baudrillard, the United States is essentially Disney World writ large. Television carries this process of abstraction to new heights, reflecting and shaping the material world in complex and highly stylized ways.

Geographies of Consumption

Drawing on the work of sociologists, historians, philosophers, and anthropologists, geographers have engaged in numerous lines of thought that suture commodities to their social and spatial origins. This body of work has tended to fall into three major categories.

First, drawing on the tradition of humanistic geography, some geographers have examined the relations between consumption, the body, and individual experience. A considerable literature,
for example, has looked at food, its origins and cultural meanings in different geographic contexts, and its role in the unfolding of daily life. Similarly, geographers have examined the shopping mall not simply as an economic phenomenon but as a cultural site pregnant with meanings. Jon Goss, for example, studied the Mall of America in Bloomington, Minnesota, which has 520 stores, chapels, a roller coaster, an aquarium, and a rain forest (see photo). In this environment, fantasy, fun, and the commodity are seamlessly merged into a seductive whole.

Second, many geographers have turned to consumption in the context of economic landscapes, including the pivotal role played by retail trade and consumer services. Traditionally, economic geography focused on production and the role of the so-called export base in economic development. When geographers turned to consumption, it was through the static and ahistorical lens of central place theory. More recent work has called attention to so-called nonbasic functions, including retail trade and personal services, and has shed light on their potential for job generation and economic change. Some geographers have studied enormous chains and franchises such as McDonald’s or Wal-Mart. Studies of the geography of tourism, which is essentially the consumption of place, are a burgeoning part of the discipline.

Third, geographers have focused on consumption in the context of the global economy, particularly the manner in which commodities are produced, distributed, and consumed via commodity chains. By embedding this sector within wider circles of finance, investment, trade, and consumption, this literature notes the ways in which globalization has unleashed a tidal wave of cheap imports that has propelled the high rates of consumer spending in societies such as the United States. This body of work traces the commodity through complex, contingent lines of causality linking sellers and buyers across multiple spatial scales. Variations of this theme point to the highly gendered nature of consumption as well as the moral and environmental dimensions that surround the way commodities are consumed, including, for example, the sacrifices made by low-wage labor trapped in sweatshops in the developing world to provide American consumers with cheap goods. Such a perspective reveals consumption as being an economic, cultural, psychological, and environmental act that simultaneously reproduces both the world’s most abstract space, the global economy, and the most intimate, the individual subject and body.

Barney Warf

See also Applied Geography; Body, Geography of; Business Geography; Commodity Chains; Cultural Landscape; Economic Geography; Food, Geography of; Geodemographics; Humanistic Geography; Identity, Geography and; New Urbanism; Retail Trade, Geography of; Sense of Place; Spaces of Representation/Representational Spaces; Sports, Geography of; Symbolism and Place; Tourism; Uneven Development; Urban Geography

Further Readings

CONTINENTAL DRIFT

See Plate Tectonics

COOK, CAPTAIN JAMES
(1728–1779)

Captain James Cook was a navigator, explorer, and cartographer whose fame is essentially due to his travels to the Pacific Ocean. His origins were humble, and he achieved his success thanks to his passion and insatiable thirst for adventure. The explorations led by him contributed enormously to the development of geographical knowledge and gave birth, in the following century, to interest in the new anthropological area of study.

After spending his childhood working in a grocery and yearning for the sea, in 1756, when the Seven Years’ War began, Cook enlisted in the Royal Navy. His map of the Saint Lawrence River played a role in the British victory in Quebec. Due to his abilities, Cook was charged with the topographic survey of the coastline and bottoms of New Scotland and Newfoundland. In 1768, Cook was sent by the Royal Society to Tahiti with the aim of having him record a rare passage of Venus across the face of the sun, but due to the lack of precise instruments, his observation could not be made in a satisfactory manner.

He then focused his attention on the exploration of the Southern Pacific Ocean to solve doubts about the supposed existence of the Terra Australis Incognita, a hypothetical continent that was imagined to be located at the southernmost part of the Earth to counterbalance the lands of the Northern Hemisphere. He reached New Zealand, mapped its coastline, and indicated the strait that separates its northern and southern islands, which have been named after him. In April 1770, Cook reached the eastern coast of Australia, where Sydney was founded.

In 1771, Cook started a new voyage. Two years later, he was the first to cross the polar Arctic Circle. He then headed toward new islands in the Southern Pacific Ocean and reached what are now known as the Cook Islands. Despite the difficult climatic conditions, he did not hesitate to challenge the dangerous polar ice and succeeded in getting as far south as 71°11′ S and circumnavigated Antarctica without sighting it. The aim of Cook’s last voyage was to get information about the existence of the Northwest Passage, a sea route through the Arctic Ocean that for centuries was supposed to be ice-free, and thus to connect the Atlantic Ocean and the Pacific Ocean. During this expedition, he mapped the coastline of Alaska and was the first European to visit the Hawaiian islands.

Susanna Servello

See also Exploration; Human Geography, History of; Royal Geographical Society
Creating a three-dimensional (3D) model of an object in a computer involves a coordinate system and algebraic calculations necessary to manipulate (change shape, scale, orientation, and location) the modeled object. Every spatial object, such as a 3D model of a house, is based on coordinates referenced to a specific coordinate system. The most popular system of coordinates used in almost every 3D modeling software is the Cartesian coordinate system, described by three coordinates, \(x\), \(y\), and \(z\), corresponding to the three dimensions of space. In the Cartesian coordinate system, all three axes are perpendicular to each other, and the system is defined by its origin \((0, 0, 0)\) and the orientation of the \(x\), \(y\), and \(z\) axes. The simplest way of referencing one object to another object is when the axes of the first object are parallel to the axes of the second object. Two other conditions still must be met. First, one needs to know the exact position between the origins of two objects. Second, one also needs to know the scale or rather the distance unit that was used in both the coordinate systems. If, for example, one house is modeled using feet and the other house is modeled using meters, then the houses will be positioned wrongly in relation to each other, and one house will be approximately three times smaller than the other house since 1 meter equals approximately 3.28 feet (Figure 1).

Projecting geographic objects such as land masses or water bodies from a 3D model of the

---

**Figure 1** Scale and distance unit implications of a coordinate system for 3D models

*Source: Author*. 

---

**Further Readings**

Earth to a 2D map is more complex than translating a 3D object from one coordinate system to another because the shape of the Earth is rather irregular compared with well-defined shapes such as a sphere or a cube. But as in the process of referencing two 3D objects, projecting geographic objects from a 3D model to a 2D map involves a coordinate system that acts as a referencing system, allowing one to translate the shape, size, and location of objects, albeit with some compromises, called distortions. Maps describe the surface of the Earth on a flat, 2D plane, which cannot be done without distortions. The distortions in shape, size, location, and, in consequence, in the angle and direction of lines, increase dramatically when the size of the mapped area exceeds a certain limit. As soon as the model grows and covers entire countries or a continent, the distortions rise to a significant problem, since the referencing between the real world and the modeled world gets rather imprecise.

There is a fairly simple and obvious solution for this problem as long as one wants to reference the real world to a virtual model, also called a model map. In a computer, there is a possibility of describing the world in all three geometric dimensions without distortions caused by moving objects from 3D to 2D representation. Global positioning systems are doing it already, and virtual globe applications are getting there. The coordinate systems that enable modeling geographic entities in 3D are called geocentric coordinate systems, because they have their origin in the center of the globe (Figure 2). Contemporary modeling software does not use
a geocentric coordinate system, but there are algorithms that can convert from one coordinate system to another.

*Erik Kjems*

See also Absolute Space; Coordinate Systems; Earth’s Coordinate Grid; Latitude; Longitude; Map Projections

Further Readings


Coordinate systems describe positions in space with one or more numbers. Coordinates can specify position as distance along a route, distance and direction from an origin, the intersection of two axes on a plane, the intersection of three axes in space, the intersection of spherical or ellipsoidal angles on the surface of the Earth, range circle or sphere intersections, a combination of vertical height and horizontal angles, or location with respect to celestial objects at a specified date and time.

Coordinate systems are sometimes called absolute methods of specifying location, as opposed to relative methods such as street addresses or proximity to landmarks. For the most part, coordinate systems can be converted, with some degree of uncertainty, from one coordinate system to another.

One of the most common coordinate systems used in geography is that of longitude, latitude, and altitude with respect to a specified horizontal and vertical datum. The spherical or ellipsoidal system of longitude and latitude angles is difficult to use for distance, direction, shape, or area estimation because longitude and latitude are not orthogonal to each other, are not equally scaled, and do not maintain a fixed relationship between their angular and distance dimensions on Earth. The Earth Centered, Earth Fixed X, Y, Z system (ECEF) defines a three-dimensional (3D) orthogonal coordinate system with the center of mass of Earth as the origin. ECEF makes vector distances and directions easy to compute in space but does not define any Earth surface and so cannot be used to compute distances and directions over an Earth surface.

To facilitate wayfinding, navigation, mapping, surveying, pointing to place, and the easy and accurate computation of distance, directions, and area, coordinate systems map portions of the Earth on a flat surface. Some coordinate systems implicitly describe point positions, while others describe areas on a surface or volumes in space. Coordinate systems are defined for local surveys, engineering drawings, regional surveying, country mapping, designation of international borders, continental and global mapping and modeling of climate change, plate tectonics, seismic activity, and exploration and mapping in space.

Surveyors and engineers often consider regions spanning less than 20 km (kilometers) to be reasonably approximated by a flat plane on which distances and directions are not appreciably influenced by the curvature of the Earth. Within 20 km, the distance difference between a line on a flat plane and the surface of an ellipsoid is less than 1 cm (centimeter). Local plane coordinate systems are defined by a point of beginning (POB), with arbitrary X and Y values, and a base line (BL) defining the rotation of the local system. Coordinates are given for other points in the system through X and Y coordinates or distances and directions with respect to the POB and BL. If the POB and BL in the local system are known with respect to another coordinate system, it is often possible to convert from a local plane coordinate system to other regional or global systems.

Small countries, provinces, states, departments, or groups of counties are often mapped and surveyed with respect to coordinate systems that are based on map projections and projection parameters. These systems allow approximate distance and direction estimates to be made with the X and Y, or easting and northing coordinates for points. More exact values of distance and direction can be computed using correction values based on the position of points and lines within the projection system. The State Plane Coordinate Systems of the United States, the British National Grid, the Costa Rican National
Grids, and the Qatar National Grid are examples. The British and Qatar national grids are based on Transverse Mercator projections, the Costa Rican grids are Lambert Conformal Conics, and the U.S. State Plane System is based on almost an equal number of Lambert Conformal Conic and Transverse Mercator grids and a single Oblique Mercator projection used for southeastern Alaska.

Larger countries, provinces, or states are mapped with projections that allow position designation but are more difficult to use for precise surveys or engineering because they are more difficult to correct over longer distances. Most of the U.S. states have statewide systems for planning and transportation management that are based on a variety of projections, including Albers Equal Area, Lambert Conformal Conics, and Transverse Mercator systems.

When printed on maps, coordinate systems can identify feature positions on or above Earth. When used in geographic information system analysis, for finding a place or navigating with global positioning system receivers, coordinates must be associated with specific units, projections, projection parameters, and both horizontal and vertical geodetic datums to be unambiguous.

An example of a coordinate system in wide use is the Universal Transverse Mercator (UTM) system. UTM was defined in the 1940s by the U.S. Army as part of the Military Grid Reference System (MGRS). UTM maps the Earth from 80° S to 84° N with a series of six-degree longitudinal strips that divide the globe into 120 mapping planes, 60 above and 60 below the equator. These zones are numbered west to east with 60 easting zone numbers and designated south to north by 20 eight-degree latitudinal northing zones. With the exception of the Central Meridian, the line of longitude at the center of each easting zone, the projection parameters for each zone are the same. Above 84° N and below 80° S, MGRS uses the Universal Polar Stereographic (UPS) system.

Within MGRS, UTM designators are given as easting zone number, northing zone character, 100 km easting character, 100 km northing character, followed by easting digits and then northing digits. The number of digits indicates the precision with which the position is given. The UTM part of MGRS is now used as the U.S. National Grid, defined with reference to the North American Datum of 1983 (NAD83).

UTM is also now used all over the world, with more than 100 UTM versions based on dozens of different geodetic datums. The UTM coordinates of a single point, often given as easting zone and northing zone, followed by the easting digits and northing digits, can differ by hundreds of meters when referenced to the wrong reference ellipsoid or geodetic datum. The same can be true for any coordinate system. Careful definition of coordinate system parameters is required for unambiguous use in positioning and analysis.

Peter H. Dana

See also Coordinate Geometry; Coordinate Transformations; Datums; Earth’s Coordinate Grid; Global Positioning System; Latitude; Longitude

Further Readings
coordinates in each reference system are used to either globally or locally calibrate a model from a set of point locations; then, the calibrated model is used to transform the data set.

Map projections are a commonly used coordinate transformation that apply a mathematical model to, for example, go from spherical coordinates of latitude and longitude, specified in degrees, minutes, and seconds, to planar map coordinates of Easting and Northing, specified in meters. This may involve a transformation between datums as well. Often, mathematical functions for converting between coordinate systems are not known. This is particularly the case for georeferencing remote sensor data or coregistering two geographic data sources to each other. In such cases, the model must be calibrated by establishing relationships between several locations in each coordinate system. These locations are referred to as control points (CP).

**Affine Transformation**

An affine transform is the most basic linear transformation model and is defined by the following equations:

\[
\begin{align*}
    u &= \beta_0 + \beta_1 x + \beta_2 y \\
    v &= \beta_3 + \beta_4 x + \beta_5 y,
\end{align*}
\]

where \((u, v)\) are coordinates in the target system, \((x, y)\) are the coordinates in the source system, and \(\beta_0\) through \(\beta_5\) are the parameters to be estimated. These equations are solved simultaneously using a least squares technique so that the root mean-squared error,

\[
\text{RMSE} = \sqrt{(u - x)^2 + (v - y)^2},
\]

is minimized. The affine transformation requires at least three CPs. Once the parameters have been estimated by solving the affine equations using CP coordinates, the transformation can be applied to the entire data set.

**Polynomial Models**

Polynomial models are often used to transform coordinates that have greater distortion or geometric error between the input and output coordinate systems that are not accounted for by the affine. The general form of polynomial models is

\[
\begin{align*}
    u &= \sum_{i=0}^{k} \sum_{j=0}^{k-i} \beta_{ij} x^i y^j \\
    v &= \sum_{i=0}^{k} \sum_{j=0}^{k-i} \alpha_{ij} x^i y^j,
\end{align*}
\]

where \(\beta_{ij}\) and \(\alpha_{ij}\) are the parameters to be estimated.

**Local Models**

In the affine and polynomial models, parameters are calibrated, then these same values are applied to all source coordinates to compute the destination coordinates. Thin-plate splines apply parameters locally across a set of data by calibrating parameters for an affine transform plus a radial basis function around CPs. Another method uses a Delaunay triangulation of the CPs and then applies a separate affine for coordinates within each triangle, where the affine is calibrated using the corner points of the triangles.

Grant Fraley

See also Coordinate Geometry; Coordinate Systems; Earth’s Coordinate Grid; Map Projections; Voronoi Diagrams

Further Readings


CORAL REEF

Coral reefs are a submarine geomorphologic feature found worldwide in tropical waters that provide habitat for thousands of species and invaluable resources for humanity. They are commonly distributed between latitudes 30 °N and 30° S. Coral reefs build in water temperatures
that range from \(~20\) to \(~30\) °C and exist in clear waters with little to no suspended sediment or dissolved organic materials. Coral reefs also tend to be located at the western edge of ocean basins. The coral organisms that build reef structures are individual invertebrates that live in symbiosis with algae organisms known as zooxanthellae. As these individual corals grow, they produce a calcium carbonate (CaCO\(_3\)) shell. Empty coral shells are the primary building material from which coral reefs are formed.

Coral Reefs: A Resource at Risk

Although coral reefs exist in nutrient-poor waters, they are one of Earth’s most diverse ecosystems in terms of species richness. Because of their high biodiversity and their efficiency in nutrient recycling, coral reefs have been likened to rain forests. Not only is the living substrate diverse, but the habitat it provides allows for thousands of vertebrate and invertebrate species to exist there. Due to the high level of biodiversity that is present within coral reef ecosystems, they provide an important resource for use by humanity. Coral reefs provide a food resource rich in protein for human consumption around the globe that includes exotic dishes for export, such as Napoleon Wrasse Lips, and local staples, such as Queen Conch. Both of these species have been fished to dangerously low levels in some parts of the world. The reefs also bring in large sums of tourism income on an annual basis for the countries where they occur. Tourists spend money in these countries on food, lodging, diving supplies, and fees to experience the coral reefs. It should be noted that

A swimmer in the Mariana Islands, Guam. Standing on or handling live corals can damage or even kill the coral polyps.

Source: David Burdick, National Oceanic and Atmospheric Administration/Department of Commerce.
even though visitors often have good intentions, they can often accidentally disturb or destroy the corals if they physically come in contact with them (first photo). Another resource that reefs provide is fauna for aquarists. The harvesting practices of aquarium fauna from coral reefs is largely unregulated and one of the most damaging. Cyanide is used by divers to stun reef fishes so that they can be captured; often the fish are evading the diver in protective crevices in the reef, and these parts of the reef are killed when they come in contact with the cyanide. The most destructive practice used to harvest stunned fish is the use of dynamite. With the use of dynamite, whole sections of coral reef structure that took hundreds of years to build are destroyed in seconds. The same reefs that provide so many resources are at risk of decline from many factors. These factors include increasing human population worldwide with increased food demands, ocean acidification, sea surface temperature increase, and harmful harvesting of fauna from the coral reef environment.

Coral Reef Formation

The geomorphology of coral reef features is formed primarily by living organisms and to a lesser extent by the cementing of available sediments. These organisms (cnidarian) provide habitat for many marine species, both vertebrates and invertebrates. There are primarily two types of corals worldwide: hard corals (scleractinians) and soft corals. The corals most responsible for reef development are hard corals, as it is these corals that build their skeletal system from CaCO₃. It is important to note that calcareous algae existing in coral reef environments also add to the structure of a coral reef.

Young coral polyps grow on top of the previous generation of deceased corals. They add to the reef structure as they build skeletal tissue. This cycle is repeated over many generations, and after a long period of time, a reef structure begins to grow. Globally, there are three main types of reef structures that are derived from this process (first proposed by Charles Darwin): fringing reefs, barrier reefs, and atolls.

A fringing reef’s development can easily be explained by using a young volcanic island as an example. In this instance, a young volcanic island is formed, creating shallow waters that ring it. If the volcanic island is in tropical clear waters, coral polyps will attach themselves to the submerged and bare volcanic rock. The cycle of growth and deposition of calcareous skeletal material begins and aggregates to form the fringing reef. If the volcanic island remains stable, the fringing reef will remain. However, if the volcanic island begins to subside, a new reef feature will result.

When a young volcanic island with a fringing reef begins to subside, a new reef structure will form called a barrier reef. After the volcano begins to subside, it will pull the coral base down with it, while at the same time, younger coral grows upward. Ocean water will break through the fringing reef and separate the fringing reef from the subsiding volcanic island with a body of water called a lagoon. Given the older age of a barrier reef, corals will be just below, and may even project above, the water surface. A good example of a barrier reef is the one that surrounds Bora Bora, French Polynesia.

If the volcanic island continues to subside, a barrier reef will become an atoll. Atolls are characterized by a coral reef encircling a lagoon, with no landmass present within the reef. The landmass has subsided below the water surface and leaves the coral reef to continue to build after it has left. An example of an atoll is the Midway Atoll, a territory of the United States.

Coral Reefs and Climate

Because of coral reefs’ persistence through time and their slow rates of accretion, they can be used to study past climates. Coral cores can provide a proxy for water temperatures, as each layer in a core represents winter and summer growth. Variations of these bands through time can give paleoclimatologists insights when deciphering past climates on Earth. Lately, environmental conditions (ocean temperature increase, acidification, and sea-level rise) on Earth have been changing rapidly. Because the rate of change is so rapid, reef-building organisms may not be able to adapt, leaving the future of coral reefs uncertain.

Jeffrey Dunn
Coral reef geomorphology is the study of the distribution, morphology, and processes that control the formation and morphological change of coral reefs and reef-associated landforms. Coral reef landforms are unique, as they result from the interaction between ecological processes responsible for the growth of coral and other carbonate producers and physical processes such as waves and currents that modulate ecological processes and redistribute carbonate material within reef systems. Coral reef geomorphology aims to understand how these process interactions vary across a range of timescales and control the morphological development of whole reef structures and reef-associated features. The range of temporal and spatial scales of interest to coral geomorphology bridges both short-term ecological and long-term geological timescales. Consequently, the discipline of coral reef geomorphology makes a major contribution to global debate regarding the future geomorphic condition of coral reefs.

Further Readings


Coral Reef Landforms

Coral reef systems possess a range of geomorphic features (Figure 1). These features reflect geomorphic development at a range of space and timescales and show varying levels of persistence in the geological record. A primary division is between *coral reefs* and *reef sedimentary landforms*. Coral reefs are three-dimensional structures consisting of veneers of living coral and reef-associated organisms that overlie sequences of previously deposited calcium carbonate separated by solutional unconformities (Figure 1). These structures evolve over geological (millennial) timescales. As geomorphic units, coral reefs range from less than 1 km² (square kilometers) in the case of smaller patch reefs to more than 100 km² in extent. Networks of reefs can form barrier complexes up to 2,400 km in length, such as the Great Barrier Reef, which is the largest biological construction on Earth.

Reef sedimentary landforms are surficial accumulations of unconsolidated sediment deposited by wave and current processes on, or adjacent to, a coral reef structure that includes reef islands and beaches (Figure 1). On geological timescales, they represent ephemeral stores of detrital material in the carbonate sediment budget. These deposits are important on the human timescale, as they form the foundation of a number of mid-ocean atoll nations, such as Tuvalu, Kiribati, and the Maldives. Both coral reefs and reef sedimentary landforms can be further divided into a suite of distinct geomorphic units that range from macro- to microscale (Figure 1).

Beginnings of Coral Reef Geomorphology

Coral reef geomorphology had its inception in the earliest modern voyages of scientific discovery to the tropical seas. Naturalists such as Joseph Banks and Alfred von Chamisso, together with navigators such as James Cook, provided the first accounts of the shape, extent, and distribution of coral reefs. These observations spawned the great *coral reef problem*, which focused on explaining the global distribution and form of coral reefs. Extensive geomorphological data sets on the apparent simplicity and recurring pattern of Pacific Ocean atolls were initially synthesized by Charles Lyell in *Principles of Geology*. These observations were later extended by Charles Darwin, whose subsidence model of coral reef development provided a
genetic sequence explaining the formation of fringing reefs, barrier reefs, and atolls.

The coral reef problem was a geomorphological problem and dominated coral reef research for more than a century. Protagonists hotly debated competing theories on the geomorphic evolution of reef structures, which included the subsidence of volcanic cones, antecedent platforms cut by waves, rising depositional banks, and sea-level change. This debate was only resolved following deep drilling expeditions, first to Funafuti, Tuvalu, and, later, to Bikini, Enewetak, and Mururoa atolls in the Central Pacific, where more than 1,000 m (meters) of shallow-water, reef limestone was recovered overlying volcanic foundations, thus confirming Darwin’s subsidence model.

**Space and Timescales of Relevance to Coral Reef Landform Development**

While coral reef geomorphology had its inception in the 19th century, largely due to the coral reef problem, it was only during the early stages of the 20th century that it emerged as a clearly defined discipline. This discipline has continued to evolve and develop a deeper understanding of the process controls on reef landform development and change. The magnitude, style, and time frames of morphological adjustment of coral reef landforms are now known to vary considerably between different geomorphic units, which in turn are dependent on the timescale and space scale of interest.

The breakthrough in explaining the broad patterning of coral reefs in the ocean basins and on
basin margins relied on plate tectonics and the global arrangement of plate boundaries. These broadscale boundary controls on reef system evolution operate at geological timescales \(10^6\) yrs. [years]. Geomorphologists subsequently recognized that on this largely geophysical template, fluctuations in sea level at timescales of \(10^3\) to \(10^4\) yrs. control the morphological expression of contemporary reefs. Oscillations in sea level, which produced alternating periods of subareal exposure during low sea-level stands and reef growth during periods of high sea level, are responsible for the contemporary morphology of individual reef platforms.

At the intrareef-platform scale, geomorphologists have resolved the influence of sea level, waves, and tidal processes in controlling the zonality of reef organisms across reefs and the formation of reef sedimentary features such as reef islands at millennial timescales \(10^2\) to \(10^3\) yrs.). At shorter timescales, coral reef geomorphologists have also examined decadal-scale changes in reef island morphology in response to climatic, wave, and sea-level variations. At the event timescale, the impact of extreme events such as tropical cyclones and tsunamis on reef structure and function have been resolved. These latter studies have contributed greatly to understanding reef island stability and change, one of the most pressing concerns of midocean nations, with the specter of future sea-level rise.

**Contributing to the Global Debate on the Future of Coral Reefs**

Due to the focus on understanding reef landform-development spanning event to millennial timescales, coral reef geomorphology makes a critical contribution to global scientific debate regarding the future trajectory of coral reef landforms. In past decades, there have been contrasting views on the future of reef systems, with geologists promoting a robust view of reefs that have survived millions of years of climatic and sea-level variation, while ecologists, based on shorter-term observations of reef ecological condition, assert that reefs are fragile ecosystems that are under threat from future global environmental change. The geomorphic timescale \(10^1\) to \(10^4\) yrs.) provides a critical link between these two temporal end members. It is able to place short-term interpretations of reef ecological condition in the context of reef dynamics over centuries and show how variations in ecology propagate through the coral reef system to influence landform change. Consequently, the discipline of coral reef geomorphology is able to address questions of landform change at timescales relevant to their management.

Paul S. Kench

See also **Atoll; Coral Reef; Geomorphology**

**Further Readings**


**CORIOLIS FORCE**

The rotation of Earth, a spinning spherical planet, has been observed to cause a change in the trajectory of any moving body or mass (e.g., a current of air or water). This change in direction is caused by the Coriolis force, a force that is in turn created by the rotating frame of reference. The Earth’s rotating frame is anticlockwise when viewing the Northern Hemisphere from space but the opposite for the Southern Hemisphere; thus, any mass given velocity across the Earth’s surface will experience an apparent force that will veer it to the right in the Northern Hemisphere and to the left in the Southern Hemisphere. Understanding the Coriolis force, named after the French engineer Gustave Gaspard de Coriolis, is fundamental to explaining major features of global atmosphere and ocean circulation, including cyclonic flow. However, the topic is often confusing, since it requires an appreciation of relative frames of reference and of scale of motion. Some essential properties of the Coriolis force are explained below.
The Coriolis force does not provide an actual push or pull on an object, nor does it originate from identifiable physical sources such as matter (as for electromagnetic or nuclear forces). Hence, it is referred to as a fictitious force. Nevertheless, within a rotating frame of reference (i.e., noninertial), the Coriolis force has real effects on motion, causing acceleration.

The Coriolis force is proportional to the mass and velocity of the moving object and acts only to change its direction, not to reduce its speed. Therefore, the apparent deflection experienced by moving objects is often referred to as the Coriolis effect.

The magnitude of this effect, or deflection, is a function of the angular velocity of rotation and the relative location or proximity to the axis of rotation. For objects in the Earth system, this means that the Coriolis force scales according to both the rotation rate (or angular velocity, $\Omega$) of Earth and latitude ($\phi$). If the surface of the Earth is considered as a disc centered at any point location of interest, then it would have a spin about the local vertical equal to $\Omega \sin \phi$. Thus, the extra-apparent Coriolis acceleration given to all object (i.e., air mass) motion around the Earth is equal to

$$(2\Omega \sin \phi)u,$$

where $u$ is the object’s speed. This means that the slower-moving winds will be less deflected than the faster ones. Similarly, a quantity called the Coriolis parameter ($f$) is defined as

$$f = 2\Omega \sin \phi.$$

The Coriolis parameter remains constant for a given latitude. Moreover, the Coriolis parameter varies between zero at the equator and maximum at poles. This means that there is no need to account for the Coriolis effect acting at the equator, whereas the deflection of motion is greatest at the highest latitudes.

The Coriolis force must be accounted for in describing the dynamics of the atmosphere and oceans that have planetary scales of motion. The effect is named after the French physicist Gaspard de Coriolis, who first analyzed the phenomenon mathematically.

**Bryan G. Mark**

**See also** Air Masses; Atmospheric Circulation; Atmospheric Energy Transfer; Atmospheric Variations in Energy; Climatology; Wind

**Further Readings**


---

**CORPORATE VOLUNTARY ENVIRONMENTAL INITIATIVES AND SELF-REGULATION**

Corporate voluntary environmental initiatives are measures that aim to improve corporate environmental performance without directly involving government legislation. Though not mutually exclusive, voluntary initiatives are considered alternatives to market-based environmental incentives, such as taxes and charges and to compliance-based ones, such as government regulations. There are several different types of voluntary initiatives, and the motivations for pursuing self-regulation vary from firm to firm.

The movement toward voluntary environmental initiatives and self-regulation came to the forefront in the 1990s. Prior to this, the 1970s and 1980s were largely dedicated to command-and-control measures by governments. Some governments now believe that firms will do the right thing on their own and therefore government interference through regulation is unnecessary (i.e., market systems and influence will prevail). The increasing deregulation of the responsibilities for environmental management has led to the development of an array of voluntary corporate environmental initiatives. In many cases, eco-efficiencies (i.e., environmental and financial savings) play a large role in determining the extent of a company’s environmental goodwill.

Some industries have developed set standards that businesses must voluntarily adhere to as a condition of membership, such as the chemical industry’s Responsible Care program. In this
way, the industry becomes, in essence, a self-regulator, relying on peer pressure and public image as motivating factors. Other forms of voluntary initiatives include eco-labeling and accreditation programs, such as ISO 14000 of the International Organization for Standardization. Environmental reporting, such as the Global Reporting Initiative (GRI), is another way in which firms can participate in voluntary initiatives (though it should be noted that reporting may not necessarily indicate an improvement in environmental performance but rather a movement toward increased transparency of a firm’s environmental performance).

Mechanisms for voluntary environmental initiatives and self-regulation can be applied separately or in combination. These range from government-to-industry pollution reduction challenges to more formal negotiated performance agreements between industry and government. In the former, government encourages the private sector to go beyond regulatory requirements, while the latter often consists of exemption from legal obligations in return for environmental improvement that goes beyond the current standards. Product stewardship programs are also gaining popularity as a type of voluntary environmental management. Liability issues are forcing businesses to rethink their environmental commitment, as banks and insurance companies are becoming increasingly stringent in their requirements of the businesses with which they work. Ideally, stewardship programs will include an examination of the entire life of a product, from the cradle to the grave. Companies with environmental management systems are also demanding that their suppliers meet similar standards. Likewise, suppliers can also influence a firm’s performance. This forces businesses to improve in order to have a competitive edge. More specifically, corporate voluntary environmental initiatives can be divided into six categories:

1. Voluntary pollution control measures by individual firms (e.g., eco-efficiencies; collective industrial ecology initiatives)
2. Government-to-industry pollution reduction challenges (e.g., nonbinding challenges initiated by government)
3. Voluntary adherence to industry sector codes of practice or other certification programs (e.g., Responsible Care, ISO 14000, eco-labeling programs)
4. Performance agreements between industry and governments (e.g., formal arrangements between industry and government in the form of regulatory exemption programs)
5. Business-to-business challenges, incentives, and agreements (e.g., product stewardship programs that involve cooperation between various supply-chain stakeholders)
6. Agreements between industry and non-government third parties (e.g., involvement of community and public interest representatives)

While a few corporate environmental initiatives are undertaken simply because they are the right things to do, most voluntary initiatives are driven by one or more perceived obligations. Motivations are based on one or more of the following:

- Financial benefits (either through long-term strategies such as investing in low-emission technologies or as short-term eco-efficiencies)
- The desire to avoid or delay regulatory action
- Fear of damage to the firm’s public image and associated customer and investor confidence
- Requirements imposed by banks/insurance companies (e.g., to avoid possible liability issues)
- Demands of suppliers or customers
- Pressure from fellow industry members
- Competitive advantage (or first-mover advantage over competitors)
- Image enhancement and/or branding
- The desire to be a good corporate citizen

The degree to which each of these incentives is important differs from sector to sector. For example, industries thought to pose a significant risk of environmental degradation can have a strong desire to regain public trust, which can be a powerful motivation for action. The financial benefits of environmental commitment range from short- and medium-term savings through improved resource management to long-term competitive advantages. While eco-efficiencies (e.g., energy
savings) may produce immediate and clearly measurable improvements, there is also a range of indirect benefits on which it is more difficult to put a precise monetary value (such as the outcomes of a firm’s improved image or increased productivity as a result of employee “pride”). Financial incentives can be further defined to incorporate money saved and money earned as the result of a firm’s level of corporate greening.

In addition to the motivations listed above, the level to which a firm participates in self-regulation or voluntary initiatives is highly dependent on other factors such as the firm’s financial position, the culture in which the firm operates, and internal leadership within the firm (e.g., champions of environmentalism). These factors are known as catalysts and can help shape a corporation’s actions by acting as a medium for encouraging or discouraging voluntary initiatives.

Voluntary initiatives are both promoted and condemned for being a replacement for regulation. Advocates of voluntary initiatives see them as being less costly and more flexible replacements for legislation. Critics see voluntary initiatives as a cover for deregulation and a mechanism for government and private sector interests to cut short-term costs by avoiding environmental responsibilities. Pure volunteerism is rare, in the sense that few firms do things to improve their environmental performance solely for altruistic reasons. Many motivations rest on the law or on possible financial gains from being green. Some would argue that corporations who put environmental and social goals over profits are in fact going against the number one goal of a firm: to make as much money as possible for its shareholders. Questions can be raised, therefore, to the extent that investors, such as shareholders, are willing to condone altruistic behavior that does not always have an immediate and positive impact on the company’s bottom line. What makes voluntary initiatives challenging is also what makes them valuable.

Jennifer K. Lynes

See also Competitive Advantage; Environmental Certification; Environmental Management; Fair Trade and Environmental Certification; Industrial Ecology; Market-Based Environmental Regulation

Further Readings


COSGROVE, DENIS (1948–2008)

Denis Cosgrove was a prolific cultural geographer who is best known for his profound impacts in the field of landscape studies. In broad terms, his work dealt with issues of subjectivity, representation, and the imbrication of power and knowledge. Cosgrove connected geographic research with the arts and humanities, providing significant insights regarding the geographic imagination and representation of knowledge, especially within the practices of landscape depiction and mapping. In 1993, Cosgrove cofounded the journal *Ecumene: A Geographical Journal of Environment, Culture, and Meaning* (now *Cultural Geographies*), a significant forum within the field. He became the inaugural Alexander von Humboldt Chair of Geography at the University of California, Los Angeles (UCLA), in 1999. Cosgrove’s contributions are associated with the “new” cultural geography, or cultural turn, and the “crisis of representation.”
Working from historical and empirical details grounded in Renaissance and Enlightenment Europe, Cosgrove’s work in cultural landscape studies focused on landscape as a “way of seeing” that represents and orders the world and humanity’s relation to it. Cosgrove placed emphasis on the ways symbolic landscapes worked to order class relations and maintain social hierarchy. Since landscape may be understood as a culturally and historically specific construction, embedded with processes of social formation and naturalizing a particular social order, Cosgrove and Stephen Daniels advocated a methodology based on iconography. Iconography seeks to uncover and interpret symbolic meanings and make visible the ideology inherent in representations. These methods have been influential, especially to geographers working in various visual forms, including feminist geographers, who have explored and modified methods based on the principles of iconography.

During the 1990s, Cosgrove shifted emphasis to explore issues of representation in mapping and cartography and to deal with broader issues of cosmology and cosmography. This period is perhaps exemplified by his analysis of images from the Apollo mission, examining the impact of visualizing the whole of Earth from space.

Cosgrove’s influence is widely acknowledged. He was an important contributor to the body of work that became known as the new cultural geography, drawing on social and cultural theory. Cosgrove and Mona Domosh helped develop the critique that became known as the crisis of geographical representation while expanding scholarly engagements with issues of authorship and authority. Through his long line of inquiry in the geographic imagination, Cosgrove’s work was informed by and helped connect diverse scholarly fields, among them art history, cartographic history, and landscape architecture.

Melinda Alexander

Further Readings


Cosmopolitanism is an ethical, moral, and political philosophy that is simultaneously very old and relatively recent. The term has multiple meanings and is often used to connote someone who is sophisticated and urbane; for this reason, cosmopolitan culture is often associated with large, internationally oriented cities. There are many varieties of cosmopolitanism, including religious and secular, and conservative and liberal.

The most common form of cosmopolitanism seeks to uncouple ethics from distance, arguing that each person is bound up with, and obligated to, humanity as a whole. Cosmopolitans are moral universalists who insist on the inherent worthiness and dignity of all individuals, irrespective of their place of birth. Morally, therefore, national and cultural boundaries are irrelevant and meaningless or, worse, distractions from the important task of caring for others in light of their shared humanity. In this view, compassion and empathy know no borders. Nationalists sometimes denounce cosmopolitans for not “being one of us.”

The origins of cosmopolitanism are often tracked to classical Greece, particularly to Diogenes, who when asked where he was from, famously replied
"I am a citizen of the world" (kosmopolite, or citizen of the cosmos), thus defying the then prevalent source of identity construction, the city-state. This notion, explicitly advanced by the Cynics and the Stoics in the 5th century BC, found its most celebrated form in Hierocles’s circle model, in which each individual is located in progressively larger webs of obligation and compassion, extending from the self and family to community to region to the world as a whole. Others find elements of this notion in the Bahá’í faith, whose founder, Mirza Husayn Ali, proclaimed, “The Earth is but one country, and mankind its citizens.”

During the Renaissance and Enlightenment, cosmopolitanism found new voices among the Western intelligentsia. Global circumvention initiated an incipient planetary consciousness among the elites of Europe, an understanding of the world as a unified entity. The 16th-century Spanish religious activist Bartolome de las Casas advocated on behalf of Native Americans and against the genocide then underway. In 1788, the German philosopher Christoph Martin Wieland wrote that all the peoples of the Earth are members of a single family. Immanuel Kant’s 1795 essay “Perpetual Peace” offers the principle that the optimal means for avoiding war was to treat all human beings as equals; he also proposed a league of nations, but his vision collapsed in the face of the advancing nation-state. After World War II, Albert Einstein, advocate of world government, dismissed nationalism as an “infantile disease.”

In this sense, cosmopolitanism is antinationalist and antipatriotic, viewing nation-states as artificial (but nonetheless very real) constructions, that is, as historical constructions, as made and not given, and therefore as mutable, plastic, and lacking inherent validity or meaning. Nationalism—with its emphasis on citizenship and on the moral geographies of inclusion and exclusion, its parochial narrowing of community to ethnic communities, its frequent xenophobia, its exaggerated importance of borders, and often its sanctification of violence against others—inhertently tends to privilege some groups over others. Without this sanctification of difference, nationalism deprives itself of the legitimacy of violence against others—in other words, war, which is its central purpose. Cosmopolitans conscientiously reject the Westphalian system of national borders as artificial and destructive constructs. As Martha Nussbaum (2002) puts it, “What is it about the national boundary that magically converts people toward whom our education is both incurious and indifferent into people to whom we have duties of mutual respect?” (p. 16). In this view, there is no such thing as an “illegal immigrant,” for no human being can be “illegal.”

In contrast, cosmopolitanism celebrates the commonalities that underlie human life everywhere, offering an “imagined community” that extends worldwide. In this respect, cosmopolitanism is not simply antiracist, antisexist, and antinationalist but is also at odds with much of the postmodern celebration of difference. Difference and diversity, to be sure, must be respected, but their assertion can also undermine the possibilities for collective action. Ironically, one of the central political projects of various postpositivist philosophies has been to uncouple difference from power. Thus, feminists attempt not to deny gender differences but to eradicate the power differentials between them, which typically foster male advantages and female disadvantages. Similarly, cosmopolitans seek to deny the power differences that accompany the accident of place of birth and national origin. Some cosmopolitanists maintain that they are “cosmopolitan patriots,” who synthesize progressive nationalism and pride of place with sympathy and empathy for others.

The growth of cosmopolitanism may be seen in light of globalization and how it has steadily problematized issues of national origin and citizenship. For some, cosmopolitanism is a neoliberal ideology, a smokescreen for those seeking to unify various nationalities under the notion of a seamlessly integrated market. However, cosmopolitanism can also serve emancipatory purposes. Just as industrial capitalism induced a “scalar shift” in identity formation from the local city-state to the nation-state, cosmopolitanism fosters a similar scalar shift from the national to the global. The moral community to which each person owes an obligation is, by definition, worldwide, generating an obligation to “care at a distance,” in which the concerns of distant strangers are held to be as important as those of people nearby. Such views are found, for example, in humanitarian interventions (by the United Nations, nongovernmental organizations, and others) and the internationalist
human rights movement in the case of genocide, famine, or other catastrophes. The ultimate expression of this line of thought is that of global citizenship, an avowed stance of many, but not all, cosmopolitans today.

Barney Warf

See also Citizenship; Identity, Geography and; Justice, Geography of; Nationalism; Social Justice; Transnationalism

Further Readings


Cost-Benefit Analysis

Cost-benefit analysis (CBA) is a decision-making support tool widely used to evaluate public sector interventions from a social perspective and in monetary terms. CBA is used to assess projects (e.g., roadway construction) as well as policies (e.g., environmental regulation). On the premise that decision making should include a comprehensive account of costs and benefits, CBA aims to provide a balance sheet of these costs and benefits.

To take the long-term impacts of a given intervention (e.g., a project) into consideration, CBA contrasts the costs and benefits of the project over its anticipated lifespan by discounting the future value of money. Essentially, the aim of CBA is to provide a net present value of the benefit-cost ratio. If this value is positive for any given project under consideration, the project is deemed to be feasible. Conversely, a negative value suggests that the project should be rejected, as society would be worse off with the project. This idea can be expressed mathematically as

$$NPV_p = \sum_{t=0}^{t} \frac{(B_t - C_t)}{(1 + r)^t},$$

where $NPV_p$ is the net present value of the project, $B_t$ are the benefits generated at time $t$, $C_t$ are the costs of the project at time $t$, and $r$ is the discount rate.

While most often used ex ante to determine the feasibility of an intervention, CBA is also used ex post to evaluate the impacts of past decisions or in medias res to evaluate a project (or policy) during its lifespan. Additionally, CBA can be employed in a comparative manner to identify one intervention of any number of alternatives that is most beneficial to society. To arrive at a recommendation, CBA generally includes the following steps:

1. Define intervention and identify alternatives
2. Identify stakeholders and objectives
3. Catalog costs and benefits
4. Predict and quantify changes
5. Monetize the costs and benefits
6. Discount costs over the anticipated lifespan

Cost-Benefit Analysis

Cost-benefit analysis (CBA) is a decision-making support tool widely used to evaluate public sector interventions from a social perspective and in monetary terms. CBA is used to assess projects (e.g., roadway construction) as well as policies (e.g., environmental regulation). On the premise that decision making should include a comprehensive account of costs and benefits, CBA aims to provide a balance sheet of these costs and benefits.

To take the long-term impacts of a given intervention (e.g., a project) into consideration, CBA contrasts the costs and benefits of the project over its anticipated lifespan by discounting the future value of money. Essentially, the aim of CBA is to provide a net present value of the benefit-cost ratio. If this value is positive for any given project under consideration, the project is deemed to be feasible. Conversely, a negative value suggests that the project should be rejected, as society would be worse off with the project. This idea can be expressed mathematically as

$$NPV_p = \sum_{t=0}^{t} \frac{(B_t - C_t)}{(1 + r)^t},$$

where $NPV_p$ is the net present value of the project, $B_t$ are the benefits generated at time $t$, $C_t$ are the costs of the project at time $t$, and $r$ is the discount rate.

While most often used ex ante to determine the feasibility of an intervention, CBA is also used ex post to evaluate the impacts of past decisions or in medias res to evaluate a project (or policy) during its lifespan. Additionally, CBA can be employed in a comparative manner to identify one intervention of any number of alternatives that is most beneficial to society. To arrive at a recommendation, CBA generally includes the following steps:

1. Define intervention and identify alternatives
2. Identify stakeholders and objectives
3. Catalog costs and benefits
4. Predict and quantify changes
5. Monetize the costs and benefits
6. Discount costs over the anticipated lifespan
7. Compute the net present value
8. Perform a sensitivity analysis

There are three criticisms regarding CBA. The first focuses on the fact that CBA considers the impacts on society as a whole rather than the impacts on individuals. While a given intervention may generate a positive net present value for all of society, the impact for individuals or a portion of society could well be negative (CBA does not take redistribution of benefits into consideration). The second criticism relates to difficulties of assigning a monetary value to certain impacts (e.g., environmental degradation) or even determining if a specific impact is considered a cost or a benefit. The third criticism is that CBA has been lacking a spatial perspective. Recently however, there have been advancements in linking CBA analysis and geographic information systems to increase the spatial sensitivity of the method.

C. Patrick Heidkamp

See also Applied Geography; Business Geography; Business Models for Geographic Information Systems; Geographic Information Systems; Spatial Decision Support Systems

Further Readings


Table 1  COMECON membership

<table>
<thead>
<tr>
<th>Full Members (Year Joined)</th>
<th>Observer Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Soviet Union (1949)</td>
<td>Afghanistan</td>
</tr>
<tr>
<td>Poland (1949)</td>
<td>Angola</td>
</tr>
<tr>
<td>Czechoslovakia (1949)</td>
<td>China</td>
</tr>
<tr>
<td>Bulgaria (1949)</td>
<td>Ethiopia</td>
</tr>
<tr>
<td>Romania (1949)</td>
<td>Laos</td>
</tr>
<tr>
<td>Hungary (1949)</td>
<td>Mozambique</td>
</tr>
<tr>
<td>Albania (1949)</td>
<td>Nicaragua</td>
</tr>
<tr>
<td>German Democratic Republic (1950)</td>
<td>North Korea</td>
</tr>
<tr>
<td>Mongolia (1972)</td>
<td>South Yemen</td>
</tr>
<tr>
<td>Vietnam (1978)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Associate Members (Year Joined)</th>
<th>Cooperation Agreements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yugoslavia (1964)</td>
<td>Finland</td>
</tr>
<tr>
<td></td>
<td>Iraq</td>
</tr>
<tr>
<td></td>
<td>Mexico</td>
</tr>
</tbody>
</table>

Table 1  COMECON membership

Source: Author.

the war-ravaged economies and their infrastructure. Invitations to participate were offered to Czechoslovakia, Hungary, and Poland, but under pressure from the Soviet Union, they declined. As a response, the Soviet Union formed the Council for Mutual Economic Assistance (COMECON). The majority of members were European, but as Soviet efforts to influence other countries gained ground, others in southeast Asia, southwest Asia, and Latin America established relationships, if not full-fledged membership (see Table 1 for membership).

**Council Session**: This was the main body of COMECON, whose decisions were not necessarily binding.

**Executive Committee**: Composed of deputy prime ministers, the group met quarterly. The body worked as the chief executive organ of the Council Session. Primarily, the Committee served to elaborate on policies and oversee their implementation between meetings of the Council Session.

**Council Committees**: These dealt with a range of issues, such as planning, metrical-technical supply,
and cooperation in scientific and technological development.

**Permanent Secretariat:** This body has been likened to a form of civil service, creating agendas for meetings and preparing statistical papers for use.

**Standing Commissions:** Several were created, focusing on issues such as the chemical industry, agricultural policy, and construction. Later in COMECON’s history, a commission on environmental protection was established.

From COMECON’s inception, multilateral trade was deliberately restricted by Joseph Stalin. After his death, the 1950s saw a trend toward increasing coordination of plans and economic integration. This resulted in attempts to foster specialization by the various members based on perceived comparative advantages. These efforts were hampered by the concept of the interested-party principle, where a member state could opt out of projects. Romania, for example, rejected the idea that it should focus on agriculture and instead sought to diversify and industrialize the domestic economy.

Bilateral trade (trade between two countries) dominated trade among COMECON members, affecting a state’s ability to seek better prices within the COMECON sphere. This was partly a function of the inconvertibility of currencies to the so-called hard currencies of Western Europe. Emphasis on domestic 5-yr. (year) plans at the expense of COMECON goals was the norm. It was claimed that trade would occur when 5-yr. plans were evaluated and countries would see what other members were producing that could meet domestic shortages.

In the absence of normal price-setting markets, barter became an important process in COMECON trade. This clearly benefited countries lacking technology as they parlayed agricultural commodities and raw materials into finished goods produced by other members. This would extend to oil and natural gas exported from the Soviet Union in the 1970s and sold at submarket prices to support other member states.

Due to problems in coordinating trade prices between states, most COMECON members practiced autarchy (efforts at self-sufficiency) at many levels, seeking to insulate their citizens from shortages stemming from missed production targets. Another issue was the tendency for the best-quality goods to be reserved for domestic consumption, resulting in shoddy, substandard goods and materials being traded with other member states.

Problems, namely, a lack of economic efficiency, plagued the organization throughout the 1970s and 1980s. Growth rates were declining throughout the period, and member states found themselves importing high-technology goods through hard-currency debt that could not be paid off easily.

Crises that affected efforts to reform COMECON included the rise of the Solidarity movement in Poland and the subsequent imposition of martial law, along with the deaths of former Soviet leaders Leonid Brezhnev and Yuri Andropov. These led to delays in talks that would address the economic issues the members were experiencing. Efforts by Mikhail Gorbachev to reform the organization and foster deeper connections were blocked by the more conservative members.

With the 1989 breach of the Austrian-Hungarian border by East Germans wishing to migrate to Western Europe, the rationale for COMECON ceased to exist. The Soviet Union did not intervene to stop the border crossings, signaling an acceptance that the COMECON members were on their way to independence from Soviet control. Eventually, the barter system gave way to trade in hard currencies, which caused a general decline in intrabloc trade. COMECON continued to hold meetings into the early 1990s. The final meeting, held in Budapest, Hungary, in 1991, saw an agreement between all members to disband the organization later that year.

Several countries expressed interest in the creation of a successor association, particularly addressing the economic concerns of Central and Eastern Europe. The demise of the organization meant that members with previously established economic specializations serving the markets in the Soviet Union lost customers as the Soviet Union reoriented its economy. Observers note that essentially, the Central and East European COMECON members traded dependence on the Soviet Union’s market for the European Union market.

*Darren Purcell*
Countermapping refers to the use of cartographic tools and maps to correct or denounce injustice. It is usually carried out in opposition to maps or spatialities produced by powerful interests, be they from the state, the private sector, or elsewhere. Countermapping’s uniqueness comes from its focus on explicit politics, advocacy, and activism. In geography, countermapping has most often been associated with questions and struggles over land use and conservation. Use of the term in the discipline began to pick up rapidly after the mid 1990s, with Nancy Peluso’s work on countermapping in Indonesian forests. Despite the recent emergence of the term in Anglophonic geography, one could trace a countermapping tradition at least as far back as the Geographical Expeditions of William Bunge in the 1970s.

Rather than a subfield in its own right, it is better to conceive of countermapping as a social practice or tactic used by many different groups dealing with a variety of issues. Several mapping tendencies are emerging under the rubric of countermapping. These include the following:

**PPGIS (public participation geographic information systems):** As the name implies, this usually consists of methods of carrying out GIS-based projects that emphasize community input and participation.

**Indigenous cartography:** This is the remapping of territories according to indigenous concepts and uses. These maps have been useful in land claims struggles vis-à-vis government authorities.

**Activist or radical cartography:** This includes the creation of maps by and for social movements to analyze problematics, form alliances, or design strategies toward a campaign. Many projects of Activist Cartography are carried out by critical artists experimenting with the tool of mapping.

These three variations of countermapping should not be understood as subfields with their own canons. Rather, they are more like areas of focus for countermapping practitioners, and each has its own trajectory and referential projects. Additionally, these three ways of engaging the practice of countermapping are not exhaustive. They are mentioned to stress the fact that countermapping is not a set of unrelated sporadic experiments at using maps to “fight the powers that be.” Countermapping in and of itself is becoming a community and network of practices, helping spread the use of maps and cartographic tools with aims of social justice.

Countermapping can include all kinds of cartographic forms, even standard ones. PPGIS, for example, often uses regular GIS representations to carry out its goals. Usually, though, countermapping also implies a critique of much of cartographic logic as such. Cartography in this sense is understood as a state-supported technology that often helps produce the situations of injustice against which countermapping is being deployed. This notion becomes clear in examples of indigenous and activist cartography. Both of these areas are developing innovative ways of presenting cartographic information, helping challenge traditional notions of the “map” or even what constitutes a map. Additionally, the tools used by countermapping projects run the full gamut of the imaginable, from the latest GIS software and locative media to sketch maps drawn by amateurs, imagination being the key limit.

Frequently, countermapping is practiced at the margins of geography departments and universities. This is especially the case with activist cartography, where small mapping groups carry out a project, often without formal cartographic
training. Countermapping is a growing phenomenon, increasingly brought up within the field of geography and increasingly used to support community struggles. With the spread of accessible geographic and mapping tools, especially through the Internet, it seems that countermapping will only spread further in the near future.

Sebastian Cobarrubias

See also Bunge, William; Cartography; Indigenous Cartographies; Justice, Geography of; Participatory Mapping; Public Participation GIS; Resistance, Geographies of; Social Justice; Social Movements

Further Readings


COUNTERURBANIZATION

The concept of counterurbanization, as opposed to suburbanization, is relatively recent in international academic literature; nevertheless, the process to which it refers is not so recent, since counterurbanization was already noticeable in 19th-century English settlements. Counterurbanization can be defined as the deconcentration of people and economic activity from urban to rural areas. It first appeared in the 1960s in the United States, giving a name to the opposite of the classic urbanization process, which involves centralizing population and economic flows in the main cities and large metropolitan areas. A process in the opposite direction began, with movement outward from large cities to smaller urban and rural settlements. This marked a step forward from traditional theories based on a hierarchical organization of settlements and the demographic and economic prominence of cities over towns and villages.

Counterurbanization represents a reversal of demographic and economic flows from larger to smaller settlements. The flows of urban population to rural areas have a varied social mix, such as retired people looking for quiet places with a healthy environment, emigrants returning to their origins after working in the city, social groups looking for different forms of life related to green philosophies (“alternative counterurbanization”), and even liberal medical, administrative, or education professionals practicing as much in public service as in private. Counterurbanization can be planned or spontaneous; regional policies may contribute to the economic development and population of rural areas, and individual people or families may change their place of residence of their own will.

France now distinguishes between different forms of demographic deconcentration processes depending on the urban or rural viewpoint; from the urban viewpoint, counterurbanization is also called exurbanization and from the rural viewpoint, rurbanization or peri-urbanization. These terms mean essentially the same thing, as they are nothing more than expansion processes away from cities caused by the spread of rapid transportation, the rejection of collective housing, anti-urban tastes in the ownership of individual properties based on green philosophies, and so on. The concept of counterurbanization is part of the transition from an industrial to a postindustrial society, bringing with it new processes for which social scientists need to define new concepts. The decline of large industrial cities and the rural resurgence bring with them a demographic
deconcentration and a change in direction from the classic country-to-city migratory flows to city-to-country flows.

Carlos Ferrás Sexto

See also Exurbs; Postindustrial Society; Suburbs and Suburbanization; Urban Hierarchy; Urban Sprawl

Further Readings


COUPLED HUMAN AND ANIMAL SYSTEMS

Coupled human and animal systems (CHAS) are those in which two elements—humans and animals (domesticated or semidomesticated)—are linked and depend directly on one another. Such systems represent one of many possible types of subsystem that exist within coupled human and natural systems. A CHAS is characterized by human and animal interactions, which create an interdependent relationship through reciprocal effects and complex feedback loops. Consequently, the system tends to react as a single entity to changes in the natural environment and to social, economic, and political human activities (Figure 1).

CHAS include various forms of nomadic pastoralism (e.g., transhumance and ranching), which, unlike farming, is an extensive form of land use typically practiced in geographical areas with little to no arable land and where farming activities would be socially and economically unviable. These locations include areas with extreme and highly variable climatic conditions, such as the arid regions of Africa and inner Asia, the Arctic and sub-Arctic regions of Eurasia, and various mountainous regions (e.g., the Caucasus, Pamirs, and Pyrenees).

Although there is much debate over what constitutes a nomadic and a pastoralist existence, one must remember that the concept of CHAS serves research-analytical purposes only. When questioning what should or should not be considered as a CHAS, rather than focusing on terminology, it would perhaps be more useful to ask oneself the following question: Can the human and animal relationship under review be considered as an integral interdependent system? If mostly it can, then CHAS would probably be a useful concept when studying how changes in the natural environment (e.g., global climate change, disease and species extinction) and social, economic, and political human activities (e.g., technology, commodity prices and trade, land encroachment and legislation) influence human populations. However, not every group practicing nomadic pastoralism should necessarily be viewed as being a CHAS. The coupled system approach would be of low analytical value for groups for which pastoralism is merely one branch of their economic mix, which, even if it has considerable symbolic or cultural value, exists alongside other branches with similar or potentially higher economic importance. In such groups, changes in the total environment (i.e., the natural environment and social, economic, and political human activities) that affect animals are unlikely to have a comparable level of influence on humans, since the latter can compensate by relying on other economic activities. The relations between humans and animals inside CHAS can be best viewed as the complex interplay that exists between animal behavior and the actions of humans. Humans’ actions are aimed at satisfying the biological needs of their animals while ensuring their own social, economic, and political needs. From humans’ viewpoint, this interplay results in achieving control over their animals to maintain their cohesion, manageability, and safety without compromising their own
requirements. For animals, the interplay with humans offers pastures and protection, which, for many species, are vital for their survival. One of the most discernible outcomes of the existing interplay within CHAS is spatial and temporal movement. Movement is achieved by humans’ skillful perception of animal behavior and their adeptness in reacting to it, which is also influenced by numerous social, economic, and political human activities. Therefore, the interplay is influenced, on the one hand, by the impact that ecological factors have on animals’ behavior and, on the other, by nonecological factors that have an impact on the actions of humans.

The utility of CHAS as an analytical concept is particularly relevant in our present-day global world. There is tremendous uncertainty and alarm concerning the impacts that anthropogenic activities have on all manner of environmental processes, ranging from the loss of large mammals due to habitat loss, to the reduction of biodiversity due to farming practices, to landscape shifts due to climate change. In addition, much uncertainty exists with regard to how economic and political change affects human groups, particularly those that are numerically small, living in discrete and relatively remote areas and having a domestic mode of economic production. Consequently, research concerned with understanding the complex systems of interaction between humans and their total environment is becoming increasingly popular. It no longer suffices to merely assess the impacts of a particular stress (e.g., a pollutant) on the environment; increasing attention is being paid to understanding the relative vulnerability of a system to a particular stress. Therefore, studying the interactions that exist between humans and animals is fundamental to assessing how external stresses are likely to influence that system.

The study of CHAS has become a multidisciplinary endeavor, largely undertaken within the field of human ecology, that includes the integration of quantitative and qualitative data. The understanding of a CHAS requires long-term participant observation of human and animal interaction, as well as biological and ecological studies aimed at understanding animal ethology. One of the greatest challenges in this endeavor is designing a common methodology.

Mark J. Dwyer
COUPLED HUMAN AND NATURAL SYSTEMS

Coupled human and natural systems (CHANS) are integrated systems in which human activities interact reciprocally with natural systems. As a species, humans’ most successful adaptation has been the ability to manipulate our environment to suit our needs. Though people have always interacted with their environment, the scope of interactions has expanded significantly since the Industrial Revolution, as population, technology, and resource use have increased. Though the term CHANS is new in the 21st century, elements of this approach have been part of geography since its inception. In particular, a focus on human-environment relations has been a major theme in geography since at least the 19th century. However, physical geographers have tended to focus on biophysical systems, and human geographers have tended to focus on social systems. CHANS adds several foci to the discussion of human-environment relations, including a systems approach, a focus on complexity, and a focus on the reciprocal nature of the relationship between humans and their environment. The interdisciplinary nature of geography has kept itself well poised to address questions regarding the interaction between humans and their environment. Unfortunately, the backlash following the rise and fall of environmental determinism caused geographers to focus on the effects of humans on the environment (environmental geography), at the expense of complex, two-way interactions that may reach from local to global scales. As a result, the varied responses of human societies to environmental change have received less attention, with the exception of hazards research. Recent projects supported by the National Science Foundation, the National Oceanic and Atmospheric Administration, and the National Institutes of Health have focused funding and attention on CHANS, not just in geography but across the sciences, making it an extremely relevant and timely topic.

Since the development of CHANS as an area of study, human-environment relations have been studied in a more holistic and reciprocal fashion to address pressing environmental challenges such as global climate change, land use and cover change, rapid urbanization, exotic species invasions, and population studies. These environmental and associated societal changes have necessitated a sophisticated and complex approach to understanding humans and their environment. Research approaches must be capable of measuring biophysical variables (land cover change, acid rain, greenhouse gases, biodiversity, and forest fragmentation), human variables (socioeconomic processes, qualitative and quantitative aspects of culture and societies, agents, policies, and governments), linkages across human and environmental spheres (resource use, natural hazards, and pollution), and complexity (nonlinearities, thresholds, feedbacks, nested scales, discontinuities, unintended consequences, emergent properties, ecological surprises). In addition, many interactions

See also Animal Geographies; Coupled Human and Natural Systems; Human Ecology; Nomadic Herding; Nomadism

Further Readings


between humans and their environment transcend single scales, so approaches must be capable of viewing problems from the scale of a single human actor up to the global scale.

The remainder of this entry is devoted to defining the major characteristics of CHANS, an exploration of climate change as a CHANS, and a review of methods and approaches used to study CHANS. The entry concludes with a discussion of recent gains in knowledge from studies of CHANS as well as limitations of CHANS for understanding environmental problems.

**Characteristics of CHANS**

CHANS are complex systems that pose unique challenges to science because of their linkages across spheres, varying spatial and temporal scales, and nonlinear behavior. CHANS, by their very nature, are defined by linkages across biophysical and socioeconomic spheres. For example, the production and transport of atmospheric pollutants, the importation and distribution of exotic species, and natural hazards all occur at the intersection of human and natural systems. Patterns that arise from these intersections and linkages cannot be understood without engaging multiple disciplines spanning the social and Earth sciences. Linkages are also characterized by reciprocal effects; human societies both act on the environment (generate pollutants) and are subject to environmental change (human health is affected by those pollutants once transported by the atmosphere). Human activities may generate feedbacks that further exacerbate (positive feedbacks) or dampen (negative feedbacks) the effects of their original activities. For instance, fire suppression in the Western United States during the mid 20th century may have led to an increase (positive feedback) in wildfire activity in later decades due to a buildup in fuel for fire. Linkages and feedbacks across spheres demand multidisciplinary science, sometimes across fields with starkly different philosophies, data, and models, so research in CHANS is interdisciplinary.

Scale is another major challenge to understanding CHANS. Interactions between humans and their environment transcend scales, including spatial scales (local to global), temporal scales (time lags from seasons to centuries), and organizational scales (individuals to international organizations). Global patterns and processes can be generated by the synergistic and cumulative effects of local processes, which repeat throughout the globe (the cumulative effect of carbon dioxide [CO$_2$] emissions by cars and power plants on the atmosphere). Relationships can develop across great distances, for example, through teleconnections in the atmosphere, through outsourcing of jobs from developed countries to India, or in the interaction between social systems and natural systems (loss of native plants due to high deer populations and export of wild harvest products). Linkages also occur across political (importation of exotic species from other countries), economic (fossil fuel burning in developed countries affecting climate and livelihoods in underdeveloped countries), and ecological (passage of migratory animals through many ecosystems) boundaries or across urban to rural gradients (solid waste transport and storage). Even the response of socioeconomic systems to environmental problems may not match the scale at which biophysical phenomena occur. For example, lands that were set aside to protect threatened or endangered species are typically too small to maintain viable populations of those species, particularly in the face of environmental change. Cross-scale phenomena make CHANS difficult to measure and manage, since methods tend to be focused at a particular spatial scale and those scales tend to be more appropriate for socioeconomic and political systems than for biophysical ones.

CHANS are also characterized by their unique properties over time. While members of Homo sapiens have always manipulated their environment to their advantage, the speed at which we have altered our environment has exponentially expanded as populations, technology, and resource use have increased. Impacts themselves have become increasingly rapid in tandem with human development. For example, globalization and rapid international transportation has led to multiple introductions of the same exotic species, yielding faster rates of spread. Since industrialization, the impacts of natural phenomena on people have become more severe as population densities in areas subject to natural hazards have increased. This has combined with the increased frequency of natural hazards in some cases. For example,
the rising frequency and intensity of hurricanes with global climate change have been coupled with increasing population densities near coastal areas, resulting in more destructive and costly disasters.

Another pattern typical of CHANS is that the effects of linkages between humans and their environment may not be detectable for years, decades, or even centuries. These long-term impacts, or legacy effects, are particularly challenging for society. Deforestation and mining in Appalachia during the early 20th century, for instance, led to long-term, lagged costs for drinking water and recreation in the late 20th century. Like legacy effects, scale mismatches over time pose challenges to policy. When the rate of change in natural systems does not match that of human institutions, societies have difficulties addressing environmental problems. One of the great challenges to global warming policy is the slow rate of temperature increase (measured in decades) relative to the term limits of elected officials (measured in years).

The unique characteristics of CHANS across spatial and temporal scales, combined with the complexity of multiple spheres (socioeconomic, political, ecological, geological, atmospheric, etc.), tend to lead to nonlinear behavior—a small change in one variable might lead to a large change in another variable. Nonlinear systems may have thresholds (transition points between alternate states or regimes) that make them difficult to predict and model. In some cases, slow processes or fine scale patterns may accumulate to produce emergent properties or surprises not immediately observable or predictable based on the individual contributions. CHANS also tend to be characterized by not one but many linkages between spheres, resulting in multivariate, interacting relationships between humans and their environment. By recognizing these nonlinear properties, it might be possible to define patterns and processes typical of CHANS regardless of geographic and cultural differences.

Climate Change as a CHANS

One of the clearest examples of a CHANS is global climate change. The way societies choose to address climate change may also serve as a model for addressing other CHANS. Since the Industrial Revolution, fossil fuel emissions, forest clearing, and biomass burning have added CO₂ to the atmosphere. This rise in CO₂ has led to increasingly serious concerns about global climate change and has generated many potential responses to mitigate and adapt to these changes. Humans interact reciprocally with the carbon cycle in a variety of ways and at a variety of spatial and temporal scales, generating a challenging system to study, model, and predict—a system that is an excellent case study of a CHANS.

While individual decisions about energy use and transportation drive individual contributions to the atmospheric carbon pool, human institutions such as private utility companies, regional and national transportation networks, and international oil and gas corporations influence CO₂ emissions regionally and globally. Direct human impacts on atmospheric carbon via fossil fuel burning are generated across several scales from individual actors (agents) up to regional, national, and international institutions. Forest clearing and biomass burning also contribute to atmospheric carbon, and these processes are similarly a function of individual decision making, local- to regional-scale economic conditions, and global economic forces. As a result, individual contributions, be they fossil fuel burning or forest harvest, have slowly accumulated to produce the emergent property of global climate change.

Human activities also have the potential to offset the rising levels of atmospheric CO₂. For example, rapidly growing forests take up large volumes of atmospheric CO₂ in their woody tissues. Local agents, such as private forest landowners, may make individual decisions about timber harvest, though these may be a function of local, regional, and global economic forces. If forces are regional or global (e.g., an international CO₂ market), then individual actions may have the potential to accumulate and influence the global carbon budget, in this case by sequestering large amounts of carbon in forests. Large industrial landowners may also influence timber harvest locally, regionally, and internationally and may respond to both local and distant economic forces. These types of local responses to distant events (teleconnections) are important characteristics of both the Earth system and CHANS.
Other important characteristics of CHANS are linkages across spheres and nonlinear changes over time. Forests in Eastern North America are currently reaching economic maturity following major cutting events at the turn of the 20th century. From a systems perspective, carbon has gradually been accumulating in these forests since about 1900, offsetting early emissions of CO₂. As these forests have matured, their rates of sequestration have slowed. Large amounts of carbon could be rapidly released when trees become large enough to mill and when markets materialize. Thus, a threshold in carbon sequestration could be reached when biophysical and economic processes align, resulting in a rapid release of accumulated carbon.

As the atmospheric concentration of CO₂ rises, local, regional, and global consequences emerge. In some cases, these consequences have the potential to feed back to human agents and institutions and ultimately to the carbon cycle itself. For example, as summers become longer and hotter, a larger proportion of homes are adding energy-intensive cooling systems—further driving energy use and additional carbon emissions. In Western North America, summers have grown longer, the duration of snow cover has been reduced, and fire activity has increased, resulting in larger CO₂ emissions from fires. Again, a positive feedback develops, resulting in major consequences for homeowners living at the urban-wildland interface, forest managers, and insurance agencies, as well as the atmosphere.

Methods and Tools

The methods and tools applied to CHANS are both quantitative and qualitative, including modeling, studies of land use and cover change (LUCC), and studies of historical land use change. Several modeling approaches have been applied to CHANS, including mathematical or statistical equations that predict human activity or environmental impacts, models that include spatial information (remotely sensed data or data collected on the ground that include spatial information) in a geographic information system, and agent-based models that focus on the behavior and interactions of individual agents. In some cases, several modeling approaches are combined, for example, agent-based models of land use. Regardless of the framework, models must be able to integrate several data sources appropriate for studying human and biophysical systems, including interviews, local and national census data, biological processes, and spatial patterns.

Studies of LUCC may test hypotheses about how CHANS influence the spatial pattern of land use or how land use and land cover may affect human decisions. Landscape patterns are measured as important emergent properties of complex systems—properties that focus attention on the interactions between humans and their environment. Studies of LUCC might ask, “Do the locations of timber mills affect the spatial pattern of forest fragmentation?” or “How does the spatial pattern of coastal settlement affect inundation following a tsunami?”

Rather than focusing on spatial pattern, another approach uses change over long periods of time to test hypotheses about CHANS. This approach integrates historical data, archaeological data, and records of past environmental change (e.g., pollen, tree rings, or sediments) to understand how different societies have interacted with their environment over time. These studies have tended to focus on the role of climatic variability in affecting the trajectory of a culture, but some have also explored how depleted natural resources can lead to the collapse or expansion of cultures. These studies provide a long-term perspective on human-environment interactions and can provide compelling stories for understanding CHANS. However, it can be argued that because of the rapid pace of current changes, modern CHANS have no analog in the past.

Theoretical Frameworks

In response to the many challenges posed by CHANS, some have adopted a systems approach, which examines the linkages and interactions between the major elements of both the Earth system and human societies. This approach is useful because it is scalable from local to global scales, applicable to many social and environmental settings, and capable of addressing some of the complex thresholds, feedbacks, and unpredictable events inherent in human-environment interactions. Using the systems approach requires
that all flows or linkages in the system are identified and quantified. This poses considerable difficulties in human systems where certain aspects of human behavior cannot be quantified or even qualitatively described effectively. The systems approach is also less a theory (capable of prediction) than a framework for examining and describing CHANS.

One metaphor capable of generating hypotheses and potentially predicting trajectories of human-environmental systems, particularly in Western management systems, is the theory of adaptive change. This metaphor is characterized by phases and cycles dependent on speeds and connections between variables. As systems move through various phases, their resilience (the capacity to retain a similar structure following disturbance) can vary, and periodic shifts from one stable state to another may be mediated by slow variables that can suddenly trigger a fast-variable response (surprise). For example, slow changes in nutrient dynamics over time, such as nitrogen fertilization from acid rain, can lead to a rapid change in ecosystem dynamics when nitrogen saturation occurs and water flowing out of a watershed is laden with excess nitrogen. The theory of adaptive change has been applied to many ecological systems that are actively managed, but it has not been applied as often in the developing world, where management of natural resources is less regulated by government agencies and is instead conducted by interacting groups of private citizens.

Future Directions and Challenges

Despite gains in recent decades, many challenges to understanding CHANS remain. Researchers must move beyond case studies and descriptions of behaviors or interactions and instead move toward theories. To do this, they must develop hypotheses that can be tested across geographic regions and environmental phenomena. Next, generalizable models will be needed; whether such models are conceptual, mathematical, or digital, they must work across scales, regions, and economic and social systems, with the ultimate goal of being able to forecast changes in resilience, ecological surprises, and responses to new perturbations. In all projects, however, researchers must recognize the uniqueness of human versus natural systems in data sources, methods, and models, taking into account the ability of humans to recognize their context, act in response to that context, and behave in sometimes unpredictable ways.

Amy E. Hessl

Gains From CHANS

Regardless of the significant methodological challenges to understanding, modeling, and predicting the behavior of CHANS, significant gains have been made in recent decades. First, it is clear that human institutions need to plan for uncertainty, consider cross-boundary effects, anticipate indirect and long-term impacts, and account for cumulative impacts in natural and human systems. By accounting for cumulative impacts, human institutions can mitigate against the loss of resilience that follows cumulative insults to a natural or human system. When systems are disturbed (e.g., by a forest fire or rapid change in a market), those systems will either return to their previous state or shift to a new stable state. The ability of a system to resist change to a new state following disturbance is a measure of its resilience. If the loss of resilience can be identified and monitored, human societies can potentially mitigate against surprises caused by rapid changes in a state that may be difficult, expensive, or impossible to reverse.

Further Readings


In geomorphology, *creep* is a noun in terms of process as well as a verb in terms of velocities. Apart from the creep of glacier ice, two kinds of creep processes are of interest to geographers: (1) the mass-movement downhill creep of rock, sediment, soil, and snow and (2) the aseismic creep, or slow movement along part of a normally earthquake-producing fault or on a volcano. In spite of their differences, both processes involve long-term and almost imperceptible slow movement of rock and earth materials. In terms of velocities of mass movement, creeping motions are described as very slow or extremely slow. Very slow rates of motion are less than $5 \times 10^{-5}$ mm/s (millimeters per second), with a typical velocity of 1.6 m/yr. (meters per year), whereas extremely slow creep rates are less than $5 \times 10^{-7}$ mm/s, or a typical velocity of 16 m/yr.

**Downhill Creep**

Creep of many different kinds of natural materials occurs as a result of constant gravitational attraction downhill. Consequentially, any slight displacements of small portions of material from many diverse causes tend to be always moved in a generally downhill direction. The net result is that in aggregate the whole mass has moved downhill with gravity.

Rock-mass creep, or deep creep, in bedrock occurs inside some mountains where many small fractures occur along foliation planes and between certain mineral grains so that spatially continuous deformation occurs. The volumes of rock masses involved are on the order of several thousands of cubic meters with thicknesses of many tens of meters. This process involves extremely slow and generally nonaccelerating differential movements among relatively intact rock masses. The deforming masses are not necessarily bounded by continuous yielding surfaces, but rather, such movements can be distributed along many internal shear surfaces that apparently are unconnected, so that the distributed shear results in bending, bulging, and folding. The total displacement is small relative to the large magnitude of the involved rock mass. The process roughly simulates viscous fluids in the distribution of velocities, with the result that some scientists refer to it as a kind of rock flow, albeit of a brittle-failure variety rather than the ductile kind characteristic of bedrock at a depth where it is under tremendous confining pressures and higher temperatures. In addition, this deep-seated gravitational creep of whole mountains can produce a sagging (*Sackung*) or settlement of the mountain so that ridge-top *grabens*, or trenches, can result, as well as *antislope scarps*, or escarpments that face uphill.

Creep of surficial sediment and soils on slopes is another of the processes characteristic of slopes worldwide, where material moves slowly downward on hillsides of low to moderate inclinations. Most of this creep is accomplished by expansion and contraction of the relatively unconsolidated materials when they freeze and thaw, or are just warmed and cooled in daily cycles with the sun, or through wetting during storms or snowmelt and later drying. In each expansion cycle, the granular soil materials move outward perpendicular to the slope, but in the return contraction part of the cycle, there is a tendency for them to move back in a slightly downhill direction because of gravity. This results in a ratcheting back and forth with a net downhill creep motion. In addition to the expansion and contraction cycles, there are also several miscellaneous additions to the creeping process. For instance, animals burrowing on slopes tend to move dirt preferentially downward with gravity, plant roots push soil preferentially downslope more than upslope as they grow, and abandoned burrows and dead
and rotted root cavities collapse downslope. Finally, deep snows on slopes push strongly downward on trees, with the result that this force is transmitted into the ground through the tree roots, with creep resulting if the ground is not frozen solid beyond the depth of the roots.

The effects of soil creep are detectable on slopes everywhere, including in tilted fences, grave stones, and telephone poles; on displaced highways and railroads; and especially in tilted retaining walls and some old house foundations. Downhill-curved tree trunks, commonly cited as evidence of soil creep, are actually almost never the result of such processes but rather occur as the result of snow creep.

The creep of snow on slopes is well-known to residents in cold climates, who commonly have to shovel out the uphill sides of chimneys, ski-lift towers, and other such structures on roofs or slopes that are in danger of being knocked over by the slow movement of the snow downhill. Snow on slopes is always creeping because metamorphism is always occurring and the porosity between grains is high. Furthermore, creep is facilitated because temperatures are normally within 95% of the melting point on the Kelvin scale. The rate of snow creep increases exponentially as the temperature goes up, especially if meltwater increases in the snow.

Aseismic Creep

Aseismic creep is a different kind of creeping motion from the types described above. It occurs as a result of the tremendous forces or pressures in the Earth resulting from plate tectonics or from volcanic activity. In these cases, the rocks move past each other slowly along surfaces of rupture or fracture without significant earthquake seismicity. Where such fracture surfaces underground are locked together with many asperities, or irregular interlocking projections, the pressure builds up until final release results in an earthquake.

In the case of strike-slip faults, such as the San Andreas in California, two enormous crustal plates slide past each other, usually resulting in abrupt and damaging earthquakes. But in some places where the fault is not locked and is waiting for a big and damaging event, it creeps slowly without any seismicity. For example, the town of Hollister, which sits astride the San Andreas Fault, is such a place and has numerous instances of slowly displaced curbstones, streets, and houses that are undergoing deformation of several millimeters (quarter to half an inch) a year, or about as fast as fingernails grow. Similarly, on some volcanoes, presumably in response to the movement of huge bodies of magma underground, fractures can result, and slow movements can occur that are also considered a kind of aseismic creep. Mount Etna, the large volcano on the island of Sicily off the coast of Italy, has fractures on it that move with aseismic creep.

John F. Shroder

See also Earthquakes; Landslide; Plate Tectonics; Rock Weathering; Soil Erosion; Volcanoes

Further Readings

Criminal Law and the Causes of Crime

Often, during the first lecture of an introductory criminology course, the professor will pose the question “What causes crime?” After a few students respond, the professor will provide the answer: “Criminal laws cause crime!” To avoid the appearance of being facetious, the professor may bring up the 19th-century European maxim of *Nullum crimen sine lege, nulla poena sine lege*, that is, there is no crime or penalty without law. Specific behaviors, processes, or arrangements that are viewed as detrimental to the common good are codified into criminal laws with sanctions, forming part of a system of formal social control known as the criminal justice system. Many criminal laws are the outcomes of political debates in which different social interests struggle to get their views enacted into policy.

Jurisprudence has vast bodies of literatures focusing on the evolution of criminal laws and their political, social, and cultural biases. There are two major schools of thought regarding the function of criminal law. The conflict school believes that criminal law and the criminal justice system are established to protect the few from the masses and that the cause of crime is external to the individual, while the consensus school believes that criminal law and the criminal justice system serve the common good and that the cause of crime is within the individual.

Approaches and Issues in Studying Crime

Crime encompasses so many behaviors emanating from different motivations and circumstances that ascertaining the geography of crime has been more suggestive than definitive. Therefore, it has been imperative to break down crime into different distinct categories and ascertain the geographic and spatial properties of the different categories. A common distinction has been to compare violent crimes (crimes against a person, e.g., homicide or rape) with crimes against property (e.g., residential burglary or automobile theft). Other approaches focus on the specific settings and places in which crimes occur, such as gang territories or drug markets.

The majority of the crime data used in contemporary research come from reports to the police, known as official reports. The geographic scale of these data range from the large scale (e.g., a burglarized home) to smaller scales representing larger areal units (e.g., census blocks, counties, or states). Another source of crime data is the U.S. National Crime Victimization Surveys (NCVS), which is intended to measure crimes not reported to the police. Currently, the geographic scales of these data are too general to yield meaningful geographic analyses. Self-reports are another source of crime information, whereby individuals are asked to give the researcher an accounting of their unreported or undetected crimes. Ethnographic studies acquire this type of data, but only a few have used any type of spatial analyses.

The metric of crime as the dependent variable is a frequency count expressed as crimes per population base. Other studies have developed different rates focused on the appropriate targets of the crime. For example, rates have been developed for sexual assaults per 1,000 females, residential burglaries per 1,000 dwelling units, and hotel/motel crimes per 100 rooms. Recently, location quotients are appearing more in the literature, providing a clearer picture of the spatial variation of crime.

The independent variables used to assess association and causation have usually relied on measures of socioeconomic status as well as housing and demographic measures. Other studies have employed variables pertaining to economic activity, city function, climate, and weather.

Back to the Future

During the 19th century, the works of the Belgian mathematician Adolphe Quetelet and the French lawyer André-Michel Guerry collectively became known as the Cartographic School of Criminology. Employing statistical methods that were novel at the time, their works examined the covariation of crime with climate, season, geographic location, gender, age, and other social and physical environment variables. Moreover, visualization of these relationships was presented in graphs and choropleth maps. Among their findings were that violent crimes are more likely to occur in warmer climates and seasons and that more property crimes occur in colder climates and seasons; the departments in Northern France have higher...
rates of property crime while those in Southern France have higher violent crime rates; males have a greater frequency and rate of involvement in crime than females; and individuals around the age of 25 have the greatest propensity for involvement in crime. Poverty and illiteracy were hypothesized as the major causes of crime; and Henry Mayhew’s study of crime in London during 1860 also found that poverty is a major cause of crime.

**The Chicago School of Social Ecology**

During the early 20th century, the Chicago School of social ecology and urban geography played a dominant role in studies of urban structure. Clifford Shaw and Henry McKay, observing the spatial distribution of delinquency and crime rates across Chicago neighborhoods, found that those with higher incidences of both were characterized by high population turnover, due to immigration, and by poverty. The landscapes of these neighborhoods contained rundown housing, factories, and abandoned buildings, with recent immigrants being the main residents. Moreover, these neighborhoods were adjacent to the central business district (CBD). Shaw and McKay observed that regardless of which immigrant group resided in these neighborhoods, the delinquency rates were always high and stable. Furthermore, they found that, as distance from the CBD increased, the crime and delinquency rates decreased (distance decay). These observations led the authors to conclude that crime and delinquency are produced by the characteristics of particular neighborhoods and that the primary causal process or mechanism is social disorganization. These neighborhoods lacked the resources and the capacity to counter the conditions creating crime and delinquency.

The Chicago School produced important benchmarks for others to follow. Their analyses of socially disorganized neighborhoods influenced subsequent researchers to examine poverty, transition, and racial-ethnic status in relation to crime. Nevertheless, the social disorganization approach became dormant because sociology-criminology changed its emphasis. There was a paucity of works replicating the social disorganization model in other cities. Moreover, some criminologists began to examine a more dynamic aspect of crime, namely, the mobility of offenders, or the journey to crime.

It is important to note that geographers did not contribute to the antecedent literature regarding the geographic properties of crime. Geographers did not start examining crime until the late 1960s.

**Geographers and Crime:**

The 1960s Through the 1980s

Crime was a major concern along with the other problems present during the 1960s and 1970s. Geographers began studying crime by publishing reviews of criminologists’ work and engaging in their own studies of the spatial variation of crime across urban areas. The advent of computers and the greater use of statistical methods meant that many of these studies were in the form of factor-ecological models, whereby a collection of variables are grouped into orthogonal factors and used as independent variables regressed against crime-dependent variables. The findings were not too dissimilar from those of Shaw and McKay and others. Social rank is negatively related to crime; measures of family status are mixed; and measures of segregation and poverty are positively related to crime. During this period, a few geographers examined the spatial behavior of robbers and juvenile offenders with respect to the distance traveled by the offenders from their residences to the crime sites. Among robbers, distances varied according to the values of their targets (i.e., the expected benefit of robbing a particular location), while the distances of juvenile trips varied with race-ethnicity, gender, and age.

During the 1980s, crime studies by geographers changed in three ways. There were fewer geographers engaged in crime research; geographers’ works began to be published in criminology and criminal justice venues; and much of the work focused on spatial behavior and interaction. For example, an ethnographic study of convicted burglars found that the locations of their crimes sites were influenced by their daily travel routines from home to work. This work integrated time budgets and routine activities to assess an offender’s spatial-temporal opportunity for crime. Another study found that rapists travel different distances to encounter victims depending on the method they use to approach their victims (e.g., a break-in
Other works incorporated the notion of mental maps and perceptions of risk and target vulnerability. The works coming from this period were less about urban structure and more about spatial behavior and interaction. They integrated time and spatial perception and presented sophisticated analyses of travel distances.

1990s Crime Mapping

During the 1990s, as the proliferation and diffusion of computer hardware and software continued unabated from the 1980s, geographic information systems (GIS) and other information systems were employed by many police agencies. Suddenly, there was a large demand for the geographer’s most faithful tool—the map. However, many of these new users of maps had very little knowledge about making or using maps. Initially, maps were used as props in operational criminal justice settings.

More widespread access to GIS, data systems, and spatial statistical software led more people to become interested in the geographic aspects of crime. Moreover, social disorganization theory was revived by the criminologists. Another group, known as environmental criminologists, grew as a result of the increased interest in crime and spatial questions.

Geographers’ contributions from the 1990s to the present have been considerable. They have been responsible for bringing some order and sophistication to the crime-mapping scene. Geographic concepts and techniques have been used to delineate and classify urban drug markets. The approach and passage of warm and cold fronts of different duration alter the expected change of the levels of calls for police services from one day to the next. Studies of spatial behavior and interaction continue. One study measuring the tendency of serial rapists to return to a crime site close to a previous one has served as one of the foundation blocks for the investigative methodology known as geographical profiling. Another study examining hot spots in different neighborhoods revealed that neighborhoods produce different crime problems with different and distinctive temporal rhythms, leading the study’s author to suggest different police responses based on the space, time, and problems.

Conclusion

What is the geography of crime? As previously stated, the answer is more suggestive than definitive. Research suggests that crime is greater in urban areas, in poorer neighborhoods, and in places occupied by disadvantaged populations. Such findings suggest that criminal laws and the criminal justice system are more focused on “street crimes,” or those in which minorities are overrepresented as victims and offenders, rather than “white-collar crimes,” which are often alleged to be comparatively underenforced, underdetected, and lightly sanctioned. The geography of crime can help illuminate such issues.

James L. LeBeau

See also Chicago School; Drugs, Geography of; Justice, Geography of; Law, Geography of

Further Readings


Crisis is an elusive term, as the controversy surrounding its causes and cures have always been unsettled among the constituents of a society. Despite its elusive nature, acknowledgment of a crisis is based on a shared view among the constituents of a system that the existing economic system cannot be reproduced without substantial reforms or displacement of the existing system. Crises thus provide society with critical junctures for progressive development as they loosen up existing negative lock-ins within the system. Recurring economic crises since the 1990s and unsettled debates over their causes and cures have presented us with opportunities to reflect on the future trajectories of capitalism as well as the substantial challenges to face with regressive aspects of the global society.

Understanding the discursive nature of the crisis diagnosis is critical as it determines the ways in which constituents respond to the crisis and address necessary steps to fix the problem. Mainstream neoclassical/institutional economics and Marxist economics differ substantially in terms of both their understanding of the crisis and their spatial and temporal implications. Mainstream economists attribute crisis to temporal failures in the localized system (e.g., moral hazards in financial markets resulting from Asian crony capitalism in the 1997 crisis) or to partial malfunction of the financial system (e.g., the credit crunch and failures of financial institutions due to defaulting of subprime loans in the U.S. housing sector). Resolution of the crisis, thus, would involve either reform of the localized system or reinforcement of supervision in the failing segments of the economic system so that natural market mechanisms could be restored. Ironically, mainstream resolution of crises has been more dependent on the workings of the state, including nationalization of failing financial firms or injection of public funds into failing financial firms, than on self-restoring market mechanisms.

In contrast, Marxist crisis theories suggest that capitalism is prone to long-term crises because of the fundamental contradictions in capital-labor relations. There have been variants of Marxist crisis theory, each of which attributes ruptures in capitalist accumulation to different causes. Among contemporaries, the overaccumulation theory of crisis (the first cut theory below) is the most widely accepted.

In The Limits to Capital, David Harvey pioneered a geographical explanation of the capitalist crisis dynamics. Harvey’s view of crises consists of three theories, which are not sequential but are rather connected via complicated dynamics. The first theory examines the ways in which the falling rate of profit à la Marx drives capitalists’ quest for surplus value–producing technologies, following the capitalist credo of “accumulation for accumulation’s sake,” and thus breaks the balance between accumulated capital and opportunities for employing capital. To fix this imbalance requires the devaluation of a portion of surplus capital and displacement of the laborer and his powers. The capitalist inner contradiction between the development of production capacity and the necessary devaluation of capital and labor, thus, supports recurring crises.

The second theory examines the ways in which the capitalist crisis tendency in production is relieved by the power of the credit system that allows investment booms in fixed capital and consumption fund formation in sales of various equities to absorb overaccumulated capital. At the same time, the power of the credit needs to be disciplined by state fiscal and monetary policies.
to ensure an expanded reproduction of capital. While the disequilibrium in production is, thus, temporarily contained by finance capital, guided through the state apparatus, new forms of instability and internal dynamics emerge as part of a new capitalist evolution.

Harvey’s geographical insight is eloquently laid out in the third theory, which integrates the geography of uneven development into the theory of crisis. The overaccumulation of the system cannot be resolved but can be relieved through localized devaluation, or spatial fix. Spatial fixes are possible because different places have different timings of upturns and downturns in accumulation cycles depending on the ways in which they are integrated into the global capitalist system. Capitals employing less competitive technologies or located in poor conditions are likely to be devalued, but the specific site of localized devaluation is hard to predict, as capitalist space economy is continuously reconfigured through territorial alliance and competition for accumulation’s sake. The crisis tendency is, thus, transformed into a pattern of compensating oscillations of local economies rather than a totalizing phenomenon.

The crisis tendency within a capitalist economy is spatially expansive as economies in crisis find temporary relief by exporting production capacity to external regions and, thus, incorporating them into the newly expanded territorial division of labor. The newly incorporated regions, however, soon reach a state where they have to seek their own spatial fixes to avoid local devaluation. This chain reaction of overaccumulation crises and spatial fixes accrues to the global crisis unless there emerges a new mode of production other than the capitalist one. In sum, space has become an integral part of the capitalist crisis dynamics through Harvey’s contribution to crisis theory.

There have been at least two additional major research efforts by geographers addressing capitalist crisis tendency and its spatial implications. First, in the 1980s and the early 1990s, geographers researched the ways in which the crisis of Fordism reconfigured uneven geographical development and was resolved through the installing of a new regime of flexible accumulation. Second, in recent research on neoliberal turns, geographers have focused on the ways in which the state apparatus has deployed neoliberal policies to resolve the long-term crisis tendency.

Crisis occupies an important role in describing uneven geographical development on Earth. In turn, the geography of uneven development facilitates the spatiotemporal resolution of overaccumulation crises through switching crises in the form of localized devaluation. This dialectical relationship between crisis and space presents geographers with ample opportunities for intervention in contemporary crisis-ridden capitalism.

Bongman Seo

See also Business Cycles and Geography; Circuits of Capital; Harvey, David; Marxism, Geography and; Restructuring; Uneven Development

Further Readings

practical, and popular. Finally, the entry indicates some of the directions in which critical geopolitics seems to be moving at present.

**Origins and Divisions**

Following the end of World War II, the term geopolitics fell out of use in the Anglo-American world, as it was seen as irretrievably connected to the academic justification of Nazi aggression. Although some political geographers continued to engage in the study of power politics at the global scale, the discourse of geopolitics lay dormant. This lasted until the 1970s, when Henry Kissinger resurrected the term as a reference to the global balance of power between the superpower blocs, thus renewing an interest in territory, space, and power that only grew with the end of détente and the beginning of the second Cold War in the 1970s.

Heavily influenced by the work of Michel Foucault, attempts to study the resurrection of geopolitics during this time period critically viewed the enterprise of geography itself as a discursive actor rather than a passive descriptor of the world it inhabits. Geopolitics, then, was an enterprise engaged in by elites (both academic and political), who attempted to remap the world in ways that were advantageous. Using Foucault’s notion of power/knowledge, it was argued that the ability to define the map on which politics would be contested was itself a key element of power. Thus, geographers (and, more specifically, practitioners of geopolitics) are key links in the construction of the taken-for-granted mappings of the world.

Less reflexively connected to the discipline of geography is the similar notion that some elites are hegemonic over the discourses of the rest of the world and, thus, have the ability to write (and rewrite) the rules by which geopolitics is conducted. Historical examples of this view include British attempts to eliminate the slave trade in the 20th century and the promulgation of a liberal individualist discourse of human rights by the United States and other actors through the late 20th and early 21st centuries. Thus, what had previously been deemed as normal and acceptable became discursively repositioned as abhorrent when the then superpower, Britain, rewrote the moral rules of slavery. However, critical geopolitics has typically been more interested in the ways in which even more elemental facets of geopolitics are naturalized. For instance, drawing from similar intellectual currents in critical (or dissident) international relations, the very primacy of the nation-state itself was questioned. A move such as this opens up vast spaces for intellectual exploration, and indeed, critical geopolitics has ranged far and wide in the years since the project was initiated. Within critical geopolitics, there has been a traditional divide of subject matter into three main categories of formal, practical, and popular geopolitics. This tripartite division is important as a reminder that geopolitics is not the study of a single set of practices and discourses but an incredibly diverse set of ideas that are quite often in conflict with one another. Thus, there is no center from which geopolitical “truth” emerges; rather, it is distilled from multiple (and numerous) centers of production (the model for this is Henri Lefebvre’s production of space).

**Formal Geopolitics**

Formal geopolitics refers to the practices and discourses associated with public intellectuals or academics in universities or think tanks such as the Heritage Foundation and Brookings Institution. This was an initial focus of much work on critical geopolitics, which cannibalized and deconstructed earlier geopolitical studies to show the workings of power/knowledge. In doing so, critical geopolitics explored the ways in which the geopolitical world, and its divisions, were produced and codified. Examples of this include Halford Mackinder’s division of the world into a “heartland” surrounded by several crescents, Nicholas Spykman’s similar heartland/rimland model, and long-standing divisions of the world into ontologically distinct oriental and occidental realms. Historically, these discursive creations use a veneer of objectivity to rhetorically advance their cause, imagining a panoptical perspective through which the world can be dispassionately observed and described. However, given the close-knit historical relationship between academic theorists and governmental apparatuses, these theories have been abundantly critiqued as being situated within nationalist perspectives that are rarely problematized. The discourses of formal
geopolitics filter through into governmental institutions through well-developed channels of journal publishing, research sponsorship, and personal relationship.

**Practical Geopolitics**

Practical geopolitics refers to the discourse and practice of politicians and governmental bureaucrats. In some ways, this is the most material of the three categories within critical geopolitics, because it occurs when the traditional apparatuses of statecraft and warfare synergize with the discourses formulated in the formal sphere. By adopting these discourses, adapting them for their own political purposes, or constructing their own geopolitical discourse in opposition to those writings, political figures promote their desired vision of the world and enforce it using the immense bureaucracy at their fingertips as well as the “bully pulpit” that comes with their positions.

Examples of this include President George W. Bush’s announcement, in 2002, of an “axis of evil” that included (and connected) Iraq, Iran, and North Korea. Regardless of any material contacts between those countries (or the lack thereof), the three were discursively connected and stigmatized as “evil” (thus implicitly leaving the United States as discursively “good”). Furthermore, the reference to the World War II Axis powers also scripted a normative response to the challenge implied in the idea of the axis of evil—that ignoring the threat amounted to appeasement. Constructions such as this are much more widely known than those of formal geopolitics because they have more immediate potential consequences, given the use of diplomatic language and practice as a way of signaling to friends and enemies alike.

**Popular Geopolitics**

Popular geopolitics is the third category of geopolitical practice and discourse. It refers to the discourse of the media and other avenues through which geopolitical knowledge circulates. While earlier practical geopolitics was described as the most material form of geopolitics because of its connection to the institutions of government, understood differently, popular geopolitics could be understood as the most material, because it serves as a sea of geopolitical narratives, scripts, and symbols in which citizens are immersed every day. This includes the direct transmission of practical geopolitical discourse (and occasionally that of formal geopolitics as well) through the traditional media, such as newspapers, magazines, and cable news networks, and also through popular culture, such as cinema, comic books, novels, television, advertising, and so on.

Popular geopolitics should not be seen as the end of the chain of discursive dissemination (i.e., from theorists to government to the public) but instead as an originator of geopolitical knowledge in its own right that in turn feeds into the other two arenas of production. For instance, many cultural elites (e.g., authors, movie directors) intentionally produce works that represent their favored geopolitical visions (or at least critique the status quo). Theorists and politicians are not immune to the influence of these cultural elites, either harnessing popular discourse for their own purposes or being legitimately influenced by it. These three categories of geopolitical knowledge (practical, formal, and popular) together constitute a continually evolving field in which the nature of the natural and taken-for-granted is contested by varying agents. Chronicling the ways in which these contests are waged, and in which certain representations emerge hegemonic over others, has been the major occupation of scholars since critical geopolitics began.

**Geopolitics in the Future**

However, while discourse remains a key part of contemporary geography both inside and outside critical geopolitics, some scholars have begun to try to push the project forward in ways that often stretch the definition of critical geopolitics. Recently, there have been two dimensions to this.

The first challenge to conventional definitions of *critical geopolitics* has come from nonrepresentational theory, which argues that geopolitics (along with the rest of geography) has been overly mesmerized by texts and representation. Indeed, as described above, critical geopolitics has been largely driven by the deconstruction of texts, whether they be old theories of geopolitics, politicians’ speeches and policy documents, or comic books and movies. These critics argue that critical
geopolitics should focus less on the cognitive elements—how people think about geopolitics—and more on the affective elements—how people feel about geopolitics. One way to do this has been to conceptualize visuality as observant practice that is itself geopolitical rather than to analyze visuality by treating images as texts to be deconstructed. Similarly, some scholars of popular geopolitics have argued for cinema to be studied not as a geopolitical narrative but rather as a vehicle for geopolitical feeling, as emphasized in the way we talk about movies—as horror, thriller, suspense, and so on.

The second challenge to critical geopolitics as traditionally constituted comes from feminist critics who claim that the focus on macroscaled global discourses neglects women, who are often relegated to “private” spaces outside the public realm of discourse of the sort studied in critical geopolitics. They argue that a more embodied, microscaled study reveals the ways in which women’s agency is brought to bear in the geopolitical realm. Furthermore, they argue that women disproportionately bear the consequences of geopolitical action, for example, when they are bombed in air raids and when they are left destitute in the aftermath of war, often without men’s assistance.

Critical geopolitics is a thriving part of the discipline but one that is evolving, with different practitioners advocating new realms of inquiry that pull it in different directions. Nevertheless, a common denominator still exists in the expectation that critical geopolitics, of whatever variety (formal, practical, popular, feminist, nonrepresentational), is intended to question the taken-for-granted definitions that dominate public debate. It has an overtly political stance that seeks to promote equitable engagement in global affairs for all.

Jason Dittmer

See also Agnew, John; Cold War, Geography of; Colonialism; Critical Human Geography; Cultural Turn; Discourse and Geography; Domino Theory; Geographical Imagination; Geopolitics; Globalization; Gregory, Derek; Hegemony; Human Rights, Geography and; Identity, Geography and; Imperialism; Justice, Geography of; Nation; Nationalism; Neocolonialism; Orientalism; Panopticon; Political Geography; Popular Culture, Geography and; Postcolonialism; Postmodernism; Poststructuralism; Production of Space; Radical Geography; Resistance, Geographies of; Spaces of Representation/Representational Spaces; Terrorism, Geography of; War, Geography of

Further Readings


Critical GIS incorporates both social theory and geographic information science (GIS) with the goal of increasing the relevance of GIS to multiple communities. The critical GIS agenda is concerned with science and technology studies, feminism, ontologies, and qualitative and participatory GIS.
While critical GIS emerged from critiques by human geographers, it has morphed in the 2000s into a creative blend of human and technical geography that has the potential to uniquely shape GIS and obliquely influence other information sciences. There are a number of issues that contribute to an understanding of critical GIS. They include its history, its relationship to science and technology studies (STS), feminism and GIS, ontology research, and public participation GIS (PPGIS).

Critical GIS did not materialize as a cohesive entity in the 2000s but rather descended from the struggles between human geographers and GIScientists in the 1990s. The 1990s witnessed a series of critiques of GIS by human geographers. That antagonism—which was influenced by the “Science Wars”—was based on concerns that GIS obeyed the dictates of the quantitative revolution. Human geography critics were keen to stress that GIS was based on Cartesian perspectivalism and, as such, could not contribute to an “emancipatory” politics. Critics were suspicious that GIS served large corporations, public agencies, and governments while eschewing the socially disenfranchised. By the end of the 1990s, however, human geography critics and GIS practitioners had reached a détente. This truce can be linked to two separate phenomena. First, GIS was pervasive by this time, and there was a general recognition that it formed a foundation of the discipline, regardless of its antecedents. Second, many GIScientists had listened closely to critics and had begun to incorporate social theory, feminism, and ontology research into their approaches to spatial research. This marked the inception of critical GIS.

This trend toward a constructive, creative critical GIS is partly due to a number of collaborative efforts between GIS researchers and social critics. At the same time, GIS researchers began to see that critical GIS is a way of democratizing the technology and extending its purview. One of the chief characteristics of critical GIS is that it is housed within the discipline rather than based on critique from outside.

Whereas external critique has little invested in the outcome, internal critique is necessarily cautious and careful, as it has a stake in the future of the technology. To be constructive, critique must actually have something invested in the subject. While critical GIS began as an external critique with an attendant harshness, it has decidedly coalesced as an influence within GIS that is wedded to the goal of a better technology. Critical GIS shoulders the burden of responsibility to a discipline while simultaneously attempting to improve on it from a theoretical and applied basis. The following section describes the fundamental research areas within critical GIS, including epistemology and ontology, feminism, qualitative methods, and public participation GIS.

**Epistemology and Ontology**

In the 1990s, critics of GIS were keen to demonstrate that GIS suffered from an impoverished epistemology that focused on quantitative data while ignoring the complexity of social systems. This is, however, not the case in the 2000s. David Demeritt, a human geographer, provides a framework for understanding the role of critical GIS in framing the epistemological impulses of GIScience more positively. He posits *heterogeneous constructivism* as a way of acknowledging that natural systems, including geographical phenomena, are influenced by a broad range of social practices but are nevertheless linked to a fundamental reality. While not explicitly recognized as a tenet of critical GIS, heterogeneous constructivism is nevertheless the epistemological rubric for much current critical GIS research. It allows researchers to demonstrate social influences in the development of technology while accepting that we can never ascertain the degree to which the science reflects phenomena as they truly are. At the same time, heterogeneous constructivism allows us to emphasize that a hybrid set of material and social processes gives rise to any technology. For instance, the development of generalization techniques in GIS was strongly influenced by a cartographic tradition that stressed map simplification rather than database generalization. Heterogeneous constructivism is a tool for studying the algorithmic basis for GIS. It implies a more GIS-savvy researcher who is interested in questions that extend beyond social construction to the better construction of GIS technologies.

While epistemology is an important factor for social theorists, GIScience researchers adopted the goals of critical GIScience in the 2000s by
focusing on ontology construction. Ontology in the philosophical sense refers to the “essence of being.” In a computing or information science context, however, it refers to the universe of discourse defined within the information system, as well as all the possible relationships between data elements. It is this latter interpretation of ontology that has given rise to the most technically oriented dimension of critical GIS.

In the information sciences, there are no foundational ontologies. Instead, each classification system, map legend, data dictionary, and semantic database constitutes a different ontology. The challenge to GIScientists is to build systems in which multiple, perhaps conflictual, ontologies describing the same set of spatial entities from different perspectives can coexist.

Feminism and Critical GIS

Critical GIS researchers have developed unique means of incorporating feminism and GIS. In the 1990s, for instance, Sarah Elwood developed a neighborhood mapping project that used GIS to study spatial problems that involve women and other disenfranchised communities. She was able to use the research to illustrate that adoption of a critical, encompassing GIS was a means of positively affecting policy at various levels of government. Sara McLafferty developed a health and GIS project that investigated environmental links with breast cancer on Long Island, in New York State. This research departed from a focus on individual behavior and cancer to one that searched for common contextual elements between women who had the disease. Each of these is a practical example of the potential to incorporate feminism into GIS.

Mei-Po Kwan has advanced the integration of GIS and feminism further by arguing that feelings and emotions can and should be incorporated into GIS analysis. Likewise, Kwan has demonstrated that it is possible to develop feminist discourses in GIS using visualization. The conclusion from this seminal work is that the technology can, in fact, be used to enrich feminist geography and practices. Furthermore, GIS needs to be reconceptualized from a strictly quantitative practice to one that embraces elements of qualitative description as well.

Incorporating Qualitative Methods Into GIScience

Critical GIS played a key role in drawing attention to the potential of incorporating qualitative data and research into spatial analysis. One of the first papers in this area was by Scott Bell and M. Reed in 2004, who argued that qualitative data could be incorporated into GIS. Since then, many papers have grappled with this issue—indicating the extent to which this dimension of research has been integrated. In 2009, Sarah Elwood and Meghan Cope published a book on qualitative GIS. This is perhaps one of the best indicators of the extent to which qualitative GIS has been accepted—and forms part of the basis for critical GIScience.

Critical GIS researchers argue that spatial events are frequently products of dynamic social processes rather than static entities. This view allows the integration of survey data with GIS to model events. Likewise, it encourages the development of analytical methods that use both qualitative and quantitative data by incorporating ethnographies. The potential of qualitative GIS is emphasized by recent trends toward social networking and information sharing via Web 2.0. Social networking applications (e.g., bird counts and other reporting of spatial/social information via the Web) have the potential to catapult qualitative spatial data into greater prominence. This outcome is closely linked to greater participation in GIS from different communities with widely varying levels of training and expertise.

Public Participation GIS

PPGIS is a key component of critical GIS. PPGIS encompasses efforts to involve diverse members of multiple communities in the use of GIS in order to inform planning and decision making. PPGIS is a flexible system of human and computer networks that integrate many methods and technologies. In addition, PPGIS incorporates multiple perspectives and a diversity of alternative information forms and facilitates collaborative planning efforts, supporting inclusive public participation in decision-making processes. Clearly, the goals of PPGIS are closely related to the emphases of an emerging critical GIS. Yet in many ways, the two niches of GIS have remained quite separate—especially as many
PPGIS experts historically published primarily in urban studies venues. Increasingly, however, the interests of the two areas have begun to converge. This is evident in the work of several key PPGIS researchers. LaDonna Knigge, for instance, used ethnographic techniques to elicit information from community members that can be integrated into a GIS framework. Rina Ghose has explored the influence of scale on access to GIS and information networks. Piotr Jankowski has demonstrated the importance of designing decision-making frameworks around the needs of local communities. Timothy Nyerges has explored the use of PPGIS by collaborative groups for community decision making. Jankowski recently extended the concepts of PPGIS to scientific decision making and demonstrated how the integration of spatial data and Web-based technologies can aid in scientific decision making. Each of these diverse examples illustrates how the concepts fundamental to critical GIS can change the everyday use of GIS when incorporated into PPGIS.

Critical GIS aims to extend the functionality of GIS while incorporating influences from science and technology studies, critical geography, feminism, ontology research, and PPGIS. As GIS researchers begin to integrate theory and technology, we can expect that critical GIS will continue to thrive. 

Nadine Schuurman

See also Collaborative GIS; Critical Human Geography; Feminist Geographies; Feminist Methodologies; GIS, History of; GIScience; Humanistic GIScience; Ontological Foundations of Geographical Data; Public Participation GIS; Qualitative Methods; Scale in GIS

Further Readings


The term critical human geography arose in the mid 1990s in Anglophonic geography as a way of representing a broad coalition of progressive approaches to the discipline. Critical human geography can be seen as a diverse set of ideas and practices linked by a shared commitment to a broadly conceived emancipatory politics, progressive social change, and the use of a range of critical sociogeographic theories. Critical human geographers draw on theoretical approaches such as anarchism, anticolonialism, critical race theory, environmentalism, feminism, Marxism, non-representational theory, post-Marxism, postcolonialism, poststructuralism, psychoanalysis, queer theory, situationism, and socialism. This entry describes the growth and development of critical human geography in primarily Anglophonic settings, and also in a number of other non-Anglophonic academic spaces (but ones that operate in, or articulate with, the wider hegemony of Anglophonic American geography). Much of the focus is on some of the key publications marking different eras in critical geography. From this perspective, the practice of critical human geography can be seen as a tentative move toward development of a historical geography of knowledge production in critical geographies.

The discussions that follow are necessarily schematic rather than a full outline of the nuanced character of these approaches. Suffice it to say that there is a good deal of overlap and a much more complex (and contradictory) character to these historical geographies of knowledge production than it is possible to present here. In outlining these various forms of geography, it is clear that critical human geography cannot be defined in a singular way but instead must be understood as multiple, dynamic, and contested and that the term critical human geographies should be used in recognition of this fact; that is, we should call them critical human geographies.

Critical human geographers are committed to transforming their worlds, and they are thus—almost by definition—engaged in significant contestation over how we should interpret existing spatial relations and how we might improve
on them. At the same time, they recognize that coalitional politics may be one of the best ways to achieve their ends.

Early Roots of Critical Geographies

The radical roots of critical geographies lie in the 19th century, among writers such as Karl Marx, Friedrich Engels, Élisée Reclus, and Peter Kropotkin. Marx and Engels are not considered geographers in the formal sense, but both contributed immensely to the critical understandings of capitalism that form the foundation for much subsequent radical geographic analyses of class relations. Moreover, some would argue that Engels’s first book, *The Condition of the Working Class in England in 1844* (published originally in German in 1845 as *Die Lage der arbeitenden Klasse in England*, then translated and published in English in 1887), considered to be a classic account of the material conditions endured by the industrial working class, can also be read as an early critical geography of urban conditions and class relations. Engels focused his analysis in part on a comparison between material conditions in cities and the countryside, and he pointed out that members of the urban proletariat in Manchester were worse off than most peasants living in the countryside.

Both Reclus and Kropotkin were anarchists who were also formally trained geographers and are understood by many to be the first radical geographers. Both are famous for writings on a wide spectrum of issues. Reclus is best known for his influential 32-volume work, *La nouvelle géographie universelle: La terre et les hommes*, for which he was awarded the prestigious Gold Medal of the Paris Historical Society in 1892 (at which time he was exiled from France for his anarchist political activism). Reclus is believed to have coined the term social geography because of his interest in the “social facts” of class struggle. Kropotkin was also a prolific writer, producing works on various aspects of the geographies of Asia and on scientific expeditions in Siberia, numerous articles in *The Geographical Journal*, and upward of 90 articles on Russian geography for the *Encyclopaedia Britannica*. He also published revolutionary pamphlets and numerous books, including two landmark treatises on anarchist praxis (*Fields, Factories and Workshops* and *The Conquest of Bread*). For critical geographers, however, Kropotkin is perhaps best remembered for his essay “What Geography Ought to Be” (first published in 1885).

A broadly Marxist tradition was continued in early-20th-century Germany by Karl Wittfogel, the first explicitly critical German geographer, with the publication in 1929 of “Geopolitik, geographischer Materialismus und Marxismus” (reprinted in *Antipode* in 1985 as “Geopolitics, Geographical Materialism and Marxism”). Wittfogel was a member of the Institute for Social Research (better known as the Frankfurt School), and like many of his colleagues, he was highly critical of the growing conservatism in Germany (and for Wittfogel specifically, in German geography). Wittfogel was an isolated critical voice in German geography; it was not until much later—in the mid 1990s—that critical geography began to rise again as an institutional force. However, as Ulrich Best points out—and in spite of the institutional resistance to critical geography—a number of student groups kept ideas for a critical German geography circulating with the production of three different, albeit short-lived, critical geography journals: *Geografiker* (1968–1972), *Roter Globus* (*Red Globe*, 1971–1973), and *Geographie in Ausbildung und Planung* (*Geography in Teaching and Planning*, 1973–1976).

Marxist traditions were perhaps more marked in France than in Germany during the 20th century. Immediately after World War II, geographers and French Communist Party members such as Pierre George and Jean Tricart wrote a number of geographical articles for the Marxist journal *La Pensée*. They had much more influence on French geography than Wittfogel did in Germany, and they began a movement that challenged the traditionalist Vidalian school. Marxist spatial analysis and geography were furthered with the development of journals such as *Espaces et Sociétés* (founded by Anatole Kopp and Henri Lefebvre in 1970) and *Hérodote* (founded by Yves Lacoste in 1976). Illustrating the ongoing relevance of anarchist and Marxist approaches in French geography, *Hérodote* has published two special issues on Élysée Reclus, in 1981 and in 2005.

Anarchism, Marxism, and socialist theory formed the foundation for the rise of what in the
late 1960s and early 1970s came to be known as radical geography in American geographic circles. Two publications of that period are indelibly linked to Anglophonic radical geography: (1) Antipode: A Radical Journal of Geography began publication from the geography department at Clark University in 1969 (first edited by Ben Wisner and then by Richard Peet from 1970 to 1985) and (2) David Harvey’s acclaimed book, Social Justice and the City, was published in 1973. Antipode continues publication to date and is celebrating its 40th anniversary. Social Justice and the City remains an important text on the reading list of almost every graduate course on the history and philosophy of Anglophonic human geography. Much of the 1970s and 1980s saw a significant output of publications and practical “interventions” in everyday life by radical geographers, most of which were dominated by anarchist, Marxist, and socialist theoretical approaches. Radical geographers were interested in using such theoretical views to expose and contest sociospatial power relations of inequity and oppression. They were also interested in better understanding the relationship between theory and practice, and they were therefore especially interested in theorizing the relationship between structure and agency. Unlike quantitative geographers, radical geographers saw human agency as important, but unlike humanistic geographers, they understood social, spatial, and economic structures to play a key role in constraining human agency. Marx’s famous dictum that “people make history, but not under the conditions of their own making” became axiomatic among radical geographers of this period.

There are other critical geographies that developed in other spaces in the 1970s and 1980s, sometimes in concert with, sometimes parallel to, and sometimes isolated from Anglo-American radical geography. Geographers in the Nordic countries began to meet in 1980 for the Nordic Symposium on Critical Geography (a tradition that continues today as a special theme in the Nordic Geographers Meetings). Mediterranean architects, geographers, and sociologists have been meeting in the Seminars of the Aegean for more than a quarter of a century to critically analyze sociospatial relations in the region. Latin American geographers, with Milton Santos (1926–2001) perhaps the most well-known among them, have also contributed significantly to international critical geographies. Finally, there is a long history of critical and radical geographies in white settler states in the Southern Hemisphere: Australia, Aotearoa/New Zealand, and South Africa. Critical geographers from all these locales have strong links with Anglo-American radical geography and wider sets of critical geographies.

The Challenge to Marxist Radical Geographies

The civil rights movements of the late 1960s led to a host of “new social movements” in the 1970s, including movements advocating for greater recognition of rights (in practice, not just in law) for people of color, women, gay and lesbian people, Aboriginal people, people with disabilities, as well as those impoverished in their role as workers (or unemployed) in the capitalist mode of production. As a result of these new social movements, radical geography branched out from its Marxist and socialist roots relatively early, with feminist, antiracist, and anticolonial works appearing in the 1970s and 1980s. Feminist geography, for example, first appeared in Antipode in 1974, with an article by Alison Hayford titled “The Geography of Women: An Historical Introduction.” Feminist articles began to appear in other journals some years later, with Janice Monk and Susan Hanson’s “On Not Excluding Half the Human in Human Geography,” for example, appearing in The Professional Geographer in 1982.

Eventually, Marxist geographies could no longer provide a single epistemological space capable of accommodating all these different forms of identity politics. To give one example, while there were many important interventions in radical geography by feminist geographers, many feminists came to believe that they were not allowed full membership in “the project” of radical geography. Accordingly, Linda McDowell wrote of being both “inside and outside ‘the project.’” The same can be said for members of other disaffected groups and their allies, including people of color, queer people, and disabled people.
Writing as disability activists (and disabled women), for example, Ali Grant and Vera Chouinard wrote of being “not even anywhere near the project.” Many of these people found intellectual inspiration in postmodern critiques of overarching theoretical metanarratives (e.g., anarchism, Marxism, and socialism) that rarely spoke to their experiences. Moreover, they engaged with postmodernism to criticize their exclusion from the project. In an important paper published in 1988, Michael Dear referred to this trend as part of a wider “postmodern challenge” to human geography. Postmodernism was always highly contested in human geography, albeit more so in the United Kingdom and the United States, where the stakes for those maintaining the status quo were much higher. Postmodernism as a set of ideas was much more welcomed in ostensibly “peripheral” spaces such as Australia and New Zealand, since geographers in these places were always both inside and outside the hegemonic project of radical geography. Notwithstanding its relatively short and contested life in geography—its currency as a theoretical approach lasted from the mid 1980s to the mid 1990s, although a closely related set of theoretical approaches known as poststructuralism continues to be a powerful force in critical geography today—postmodernism played an important role in a number of transformations of the discipline. One significant impact that postmodernism had on human geography was to precipitate a crisis of representation that in many respects continues to affect critical geographies to this day. Partly as a result of this crisis of representation and partly as a result of not personally identifying as Marxist (although all would identify as Left-progressives), very few geographers now use the label radical geography to define their work. Instead, most use the broader term critical geography.

The Critical Geography Forum and the International Critical Geography Group

The term critical geography had begun to be used by geographers in the 1980s and 1990s. The Nordic Symposium on Critical Geography, for example, was initiated in 1980. The Cincinnati Mini-Conference in Critical Geography started in 1994. Macmillan Publishers even began to market a book series under the name “Critical Human Geography” in 1982. However, critical geography became institutionalized in geography as a term generally accepted by a broad constituency of Anglophonic geographers only following a series of events arising from the merger of the Institute of British Geographers (the scholarly society for academic geographers) with the much larger Royal Geographical Society. The story is worth recounting here, as many of the concerns raised are emblematic of the way critical geographers tend to operate as intellectuals, as knowledge workers, and as citizens of the world. The merger formed what became the Royal Geographical Society (with the Institute of British Geographers) (RGS-IBG), and it confronted Left-oriented geographers with a series of political and ethical dilemmas arising from the sponsorship of the RGS-IBG by the multinational company Royal Dutch Shell. First, many geographers had concerns about the RGS-IBG being sponsored by multinational capital, with its attendant exploitation of workers (through the appropriation of the surplus value of labor); they were also concerned with issues of indigenous rights, environmental justice, and environmental racism, and they were concerned with the neocolonial relationship that the RGS-IBG was implicated in because of its acceptance of money from Shell.

Shell had for some years prior to the merger of the RGS-IBG been engaged in oil extraction activities in the Niger River Delta in Nigeria. The Ogoni people, who lived in the delta, had long criticized Shell’s activities in the region—complaining in particular about its devastation of the environment and its lack of concern for local peoples and also about Shell’s indirect support of the military junta that ruled Nigeria through oil revenues accruing to the state from Shell. On November 10, 1995, nine Ogoni people, including the famous writer and political activist Ken Saro-Wiwa, were executed by the military regime. According to a press release dated April 6, 2009, Shell was to stand trial in a federal court in New York State on charges that it was complicit in human rights abuses in Nigeria, including the execution of the nine Ogoni people, as well as in crimes against humanity, torture, arbitrary arrest, and detainment.
As a number of contributors to a special issue of *Ethics, Place and Environment* noted, Shell’s activities in the Niger Delta and its links to the Nigerian military regime posed significant problems for the RGS-IBG, given that Shell was a major sponsor of the organization. After the execution of Saro-Wiwa and his colleagues, a campaign was launched to convince the RGS-IBG to end its relationship with Shell. Just prior to the execution, a number of critical geographers in the United Kingdom had already begun to explore options for developing an alternative geographical organization (of some sort) to the RGS-IBG, and some of these geographers created the Critical Geography Forum online (Joe Painter is the person credited with actually creating the list on the U.K.-based system in 1995). The Critical Geography Forum list came into its own as the space for discussions in support of the campaign to end Shell’s sponsorship of the organization. Academic geographers voted overwhelmingly in support of a motion to end the sponsorship, but this was ignored by the RGS-IBG Council. Eventually, the critical geographers were successful in getting a motion to end the sponsorship put to a membership vote at a special general meeting of the RGS-IBG in November 1995. That vote was dominated by former RGS members (nonacademics, as opposed to the academics who made up the former IBG), and the motion to end the sponsorship was overwhelmingly defeated.

While the motion to end Shell’s sponsorship of the IBG was defeated, the Critical Geography Forum became a venue for discussions that led to creation of an alternative geographic organization and for organizing alternative conferences. The inaugural International Conference of Critical Geographers (ICCG), which took place in 1997 in Vancouver, Canada, was organized by members of the Critical Geography Forum and gave rise to the International Critical Geography Group (http://econgeog.misc.hit-u.ac.jp/icgg). Subsequently, the International Critical Geography Group organized four more International Conferences of Critical Geography (in Taegu, Korea, 2000; Békéscsaba, Hungary, 2002; México City, Mexico, 2005; and Mumbai, India, 2007).

Critical geography is now institutionalized in much of Western academe. It is arguably hegemonic in places such as the United Kingdom, and it is certainly ascendant in places such as Australia, Canada, Denmark, Finland, New Zealand, Sweden, and the United States. Critical geographers often now hold senior positions in both their universities and the wider academy in many of these places. This positioning of critical geography as almost normative has transformed human geography, but it has also had transformative effects on critical geography itself. As part of a now “mature” approach in human geography, critical geographers are much less likely to be making programmatic statements about what geography ought to be. Instead, critical geographers are using a wide range of critical social theories to undertake substantive studies of people’s geographies. At the same time, critical geographers are subjecting various forms of critical geography itself to interrogation for their contribution to classism, colonialism, inequality, injustice, racism, and sexism. Finally, many critical geographers (often under the rubric of nonrepresentational theory and/or theories of practice) are engaging in work that radically questions how we know our world and how we represent it, which might possibly bring us back again to an all-new kind of radical geography.

Lawrence D. Berg

See also Anarchism and Geography; Antiglobalization; Antisystemic Movements; Bunge, William; Critical Geopolitics; Cultural Turn; Disability, Geography of; Environmental Justice; Environmental Racism; Feminist Geographies; Gays and Lesbians, Geography and/of; Gender and Geography; Harvey, David; Human Geography, History of; Identity, Geography and; Imperialism; Justice, Geography of; Marxism, Geography and; Masculinities and Geography; Massey, Doreen; Mitchell, Don; Orientalism; Other/Otherness; Peet, Richard; Political Ecology; Political Economy; Postcolonialism; Postmodernity; Poststructuralism; Queer Theory; Race and Racism; Radical Geography; Smith, Neil; Social Geography; Social Justice; Walker, Richard; Whiteness
The different ways in which people understand nature have been at the heart of struggles over what kind of discipline geography ought to be. Recent years have seen the emergence of a range of critical approaches to the study of nature that seek to critique and change it. Nevertheless, conceptually, nature remains notoriously difficult to define, and as will become clear, definitions often reveal much more about an overall political position—and a view about what geography ought to be—than they do about plants, animals, landscapes, and the many things we think of as nature.

In its earliest manifestations, geography was defined by an emphasis on the manner in which natural environments shaped and determined human societies. Here, nature assumed a determining role over the social: Differing levels of civilization were interpreted through this deterministic lens. This position was brought into question by those who sought to place a greater influence on the role of culture as a mediating factor between the social and the natural. Nature, in the process, came to be understood as something imbued with a variety of social and cultural characteristics. Rather than being seen as a pristine realm, devoid of human influence, nature was viewed as a product or construct. These constructions of nature have varied historically and geographically. Subsequent debates have followed similar lines to these first confrontations, with contemporary critical work increasingly assuming an antiessentialist or denaturalizing track. To put things crudely, the aim has been to show that nature is anything but natural.

As several authors have noted, the dominant view of nature has served to deepen the gulf between human and physical geography. However, more recent critiques have sought to establish the myriad ways in which apparently discrete social and natural entities are actually co-constitutive. Although in according nature a greater agency this might seem to be a move back toward the ground of environmental determinism, the thrust of this work is entirely different. Instead, the aim has been to dismantle the false dichotomies through which nature and society have been read and interpreted. Perhaps here, there may be more grounds for a conversation, and it is perhaps no surprise that environmental geographies seem to be undergoing a minor renaissance. In what follows, this shifting focus from the social construction of nature toward a growing recognition of the mutually constitutive relationships between nature and society is charted. At the same time, the important synergies with other approaches, from feminism to Marxism, are captured.

Constructing Nature

If the overarching thrust of critical work has been to undermine the authenticity of the natural, this has tended to be through a focus on historically and geographically varying constructions or representations of nature. Different meanings are shaped by dominant representations of nature in
a particular place at a particular time, as well as an individual’s place within a particular social structure. In spite of these particularities, it is possible to isolate three common understandings. Thus, nature is (1) the nonhuman world (and occasionally the human and nonhuman worlds together), (2) an inherent force directing the human or nonhuman world, and (3) the essence of something (as in a thing’s “natural properties”). These definitions are potentially contradictory, again merely reemphasizing the real difficulty in isolating what exactly nature is. Critical studies of nature have picked up on these ambiguities while demonstrating the way in which meaning is constructed through particular relations of power. Here, the apparent neutrality of representations of nature has been exposed to be ideologically driven. Landscape painting, as one dominant mode of representing the beauty of a particular environment, has been presented as a representational practice that enables everything from property rights to gender relations to appear as natural or authentic expressions of the way the world ought to be. Both property rights and gender are social constructions, but associating them with the natural world appears to give them a much stronger legitimacy. They come to be dominated by a force outside human control. Because of this, work on the social construction of nature has had a broadly liberatory agenda. It seeks to demonstrate how many of those characteristics considered to be natural are social constructions. Rather than contenting ourselves with the world as a natural given, the possibilities for a better, more egalitarian, and more ecologically sustainable future begin to emerge.

Political ecologists have worked with similar tools to demonstrate the ways in which environmental problems cannot be understood outside the power relations, institutions, and economic processes of which they are a part. Some have referred to the focus of political ecological research as “the politicized environment,” and one of the key contributions has indeed been to politicize understandings of nature. To take a classic example from the literature, soil erosion in Nepal was for many years linked to both geomorphological processes and the specific agricultural practices of farmers. Combating soil erosion required a change in cultural practices and a welter of development praxis focused on reeducating local farmers to care for their lands better. For political ecologists, it is vital to situate the decisions being made by farmers within broader systems of accumulation. Soil erosion, in short, should be seen to emerge from the limited options available to small-scale farmers within a capitalist society. Geomorphological processes and indeed the soil itself are crisscrossed with broader relations of power.

Turning their attention to famine, political ecologists have also played a key role in undermining a resurgent neo-Malthusianism in the late 20th century. Through in-depth ethnographic research, they have exposed the social causes of what some interpreted as natural disasters. Michael Watts, for example, has demonstrated that West African farmers coped successfully with drought until colonialism undermined indigenous food-producing systems. Here, there are connections with hazards research, which has similarly been an important terrain for pioneering critical approaches to understandings of nature. As with work in political ecology, critical hazards research has moved away from explanations of disasters based on natural processes and individual decision making. Instead, decisions have been rooted within broader political economy questions and linked more fully to the systems of accumulation in which hazards emerge.

Denaturalizing Gender and Socializing Nature

Feminist geographical research has similarly had a central role in shaping critical studies of nature. From the 1970s onward, feminist geographers have exposed the many ways in which apparently natural differences between the sexes have been produced and reproduced through a variety of social codes, representations, and practices. As with the question of nature, dominant epistemological frameworks have posited a series of dualisms around which differences are reinforced. Above all, femininity has been associated with nature, whereas masculinity has come to represent...
the domain of culture. Again, these dualisms have been reinforced within literary and artistic forms of representation, as women have been visually and metaphorically connected to “Mother Earth.” These tropes are mutually reinforcing. Thus, a feminized view of nature emerges at the same time as a naturalized view of femininity. Exposing these dualisms and demonstrating their subversive operation in varying historical and geographical contexts has been the central aim of much feminist research. Paradoxically, certain approaches to researching gender and nature appear ultimately antithetical to this denaturalizing impulse. Biological ecofeminism has celebrated women’s proximity to nature as opening up possibilities for a new ecological consciousness. Although such a biologically determinist position is rejected by the vast majority of both feminist environmentalists and also ecofeminists, it nevertheless reopens aspects of the debate over nature. For more socially minded ecofeminists, divisions between men and women are the result of a gendered division of labor (and thereby a product of society and not nature). Thus, an interpretive response to a structurally determined (and socially produced) position is required if a radical ecological consciousness, a transformed conception of nature, and a challenge to patriarchy are to emerge from women’s role within broader ecosystems.

More recent work in feminist geography has sought to dismantle the assumption that gender is a social construct while sex and the body are natural givens. Again, the tendency is to naturalize what is actually a more complex social product. Bodies, in short, should be seen to be actively coded and produced. Neither fully natural nor fully social, they are a product of mutually constitutive sociocultural relationships. Perhaps some women do indeed “throw like a girl,” but this cannot be seen outside the particular functions, movements, and performances that bodies are socially expected to undertake. Once again, this blurring of social and natural categories has both mirrored and informed critical studies of nature.

**After Humanism**

Integrating some of these insights more directly into critical studies of nature, researchers have sought to show the manner in which nature is enacted or performed. This has emerged most clearly within work that is often referred to as post-humanist. Criticizing humanist thought for a hubristic overestimation of the agency of humans, post-humanists have sought to develop an agenda for researching interspecies connections. Whereas the rigid compartmentalization of natural and social processes might be seen to emerge from humanist perspectives, the notion that nature is a social construction is also criticized. The epistemological critique of the latter has been replaced by an ontological critique. In short, dividing the world into boxes labeled nature and society is seen to make little sense in reality, not to mention our ways of thinking about reality. Seeking to capture this reality within a more appropriate language, theorists have described a variety of socionatures, cyborgs, hybrids, and imbroglios.

Although now common within geographical work, most post-humanist writing has emerged from social studies of science and has been influenced in large part by French social theory. Perhaps the most common approach within such work has been actor-network theory. More clearly defined as an approach to social scientific research than as a coherent theoretical framework, actor-network theory has sought to demonstrate the manner in which reality is made up of networks of human and nonhuman actants (actants are essentially actors but can include nonhuman objects, which in actor-network theory enable and constrain the actions of others and thus have consequences). The term actants refers to the fact that neither the human nor the nonhuman is genuinely an actor. Instead they both depend on the other actants that constitute these broader networks. From studies of scallop fishing, in which the scallops themselves were viewed as actants in a chain of processes surrounding their harvest, to research on the practice of scientific research, this work has done much to trace the numerous connections that make up reality. Clearly, it becomes difficult to sustain any rigid boundaries between nature and society or between humans and nonhumans, as networks appear to defy any socially imposed framework of understanding. The division of geography into its social and natural scientific branches becomes increasingly problematic: Any discussion of a “human geography” is to automatically short-circuit a much broader network of
connections. More recently, geographers have been drawn to hybrid geographies and explorations of the myriad mixings between humans and nonhumans as ways to avoid a strict dichotomy between the social and natural worlds and emphasize that they are inescapably entwined.

Clearly, the post-human actor-network theory approach has appealed to those broadly interested in an ecosystemic approach to nature, and for some people, there are important connections to deep-green or deep-ecology philosophies. As with some deep-ecology work, the post-humanist emphasis on nonhuman agency has been criticized for its inability to address socially produced power differences, its inattentiveness to broader political economy questions, and its problematic approach to history. It can also exhibit a naive celebration of all life forms regardless of their impacts on human beings. For example, there is little to celebrate about the virus that causes AIDS or plague-carrying rats. Linguistically, there are also problems—the very term *hybrid* suggests a meeting of two things (whether species, inanimate objects, or whatever) that were formerly distinct. The point of post-humanist approaches such as actor-network theory is to overcome deeply embedded dichotomies such as between humans and nature and to stress their interconnections in such a way as to emphasize their inseparability. This approach has led to a blurring with more populist theses on the so-called end of nature, the notion that the natural world has been so utterly transformed by human beings that it is impossible to discuss it without reference to how people have constructed, modified, represented, and reproduced natural systems of energy, matter, and life. Here, a new era of techno-natural hybrids has been either celebrated or condemned. Even if the thrust of post-humanist work is to argue that we have never been fully human, it has been difficult to find a language that does not reproduce the sense in which humans and nature are only now beginning to intermingle. Nevertheless, from the turn of the millennium, research in hybrid geographies seems to have captured the imagination of many in the field of Anglophonic geography, inspiring a wealth of studies across the discipline and also bringing about some fascinating collaborations between human and physical geographers.

### Producing Natures

Many of the criticisms of hybrid geographies have come from Marxist scholars, who claim that the inability to consider capitalist relations makes such research academically defunct and politically impotent. The result has been a muted standoff between two of the most important contributions to critical studies of nature within geography. The potential overlaps and grounds for dialogue have thus been neglected. Outside political ecological work, the key contribution of Marxist approaches to critical studies of nature has been to understand the latter as produced. Although this has echoes of the social constructionist position, it differs in several significant ways. First, as with post-humanist critiques, it develops an ontological argument against the separation of nature and society. Second, and relatively, it focuses on everyday practices and, above all, on the labor process rather than on representations and ways of thinking about nature. For Karl Marx, work or practical activity is a metabolic process. By this, he means that through the creative act, nature is transformed at the same time as humans are transformed. Indeed, they are indissociable. If a carpenter produces a chair, this not only transforms the wood being worked on, it also provides something of use for society that changes the way it is organized. Although this applies directly in societies in which production is for the immediate satisfaction of one’s own needs or the needs of one’s direct friends or family, it becomes more complicated when commodities are produced for exchange. For exchange to operate effectively, a universal measure of value is needed—something that transforms the myriad qualities of a commodity (e.g., a chair’s ability to provide rest, comfort, or relaxation) into a single, universally recognized measure of value. Money performs this role, and with the entry of money into exchange societies, nature is increasingly recognized for both exchange value and use value. The work process thereby serves to mediate the transformation of nature into both a use and an exchange value. Capitalist societies imbue a further set of relations in this production of nature, and nature comes to embody the relationship between worker and capitalist. Important within this framework is a sense that nature has always
been produced. It is clearly possible for humans to relate to one another in ways that are more or less beneficial for other life forms; however, the interdependence of ecosystems remains transhistorical. These insights have spawned work on cities as particular created ecosystems or, as some have referred to it, on the urbanization of nature.

If this historical materialist approach to the study of nature shares several of the insights of hybrid or actor-network approaches, it is in the sense that, ontologically, nature and society are one. For Marxists, the language adopted has tended to emphasize the manner in which nature is and has always been a differentiated unity comprising the social and the natural. For some, this is best referred to as a dialectical relationship. With the focus on historically and geographically specific relations, research on the production of nature has foregrounded the importance of capitalist forms of organization for questions about nature. Indeed, for historical materialists, none of the key questions around nature can be adequately understood unless some weight is given to these social relations. Malthusian approaches to nature and resources, for example, are turned on their heads if the necessity for capitalism to produce a surplus population is recognized. Resources come to be relationally defined. If those relations change, so do the resources and the key questions being asked around them. Here, there are connections back to research in political ecology, which has also been open to using the methodological tools provided by Marx as a form of nature critique. Finally, critical work on nature from a Marxist perspective has developed interesting connections with work on environmental justice. If nature is actively produced through the work process, it is the environment lived and experienced by people on an everyday basis. Rather than defending wilderness areas, environmental justice is about challenging the injustices that result from some people being forced to experience a dirty, polluted nature that harms the health of themselves and their families on a daily basis.

Conclusion

If the concept of nature employed in critical human geography appears to be some distance from that in the popular imagination, this is partly an intentional move. Geographers have sought to expose the ways in which certain concepts of nature can have a paralyzing influence on a politics of liberation. At the same time, the example of environmental justice shows the way in which it might be possible to reconstruct a politics of nature that is critically relevant to everyday life. Overall, critical studies of nature appear to have moved from a strong social critique of the very concept of nature to a more qualified recognition that nature matters. An epistemological critique of the separation of nature and society has given way to an ontological argument about internally related socionatures. While these debates have inevitably been shaped by academic fashions, they have real import for both the type of discipline geography might be and the makeup of the world we wish to live in.

Alex Loftus

See also: Actor-Network Theory; Class, Nature and; Critical Human Geography; Cultural Ecology; Ecofeminism; Ethnicity and Nature; Gender and Nature; Marxism, Geography and; Nature; Nature-Society Theory; Political Ecology; Race and Nature; Social Construction of Nature

Further Readings


Crop genetic diversity describes the genetic diversity within or between plant species that are either directly useful to humans in the context of food and agriculture or are relatives of such species.
To increase the genetic diversity of U.S. corn, the Germplasm Enhancement for Maize project seeks to combine exotic germplasm, such as this unusually colored and shaped maize from Latin America, with domestic corn lines.

Source: Keith Weller/ARS/USDA.

This genetic diversity can be assessed at different taxonomic levels, such as families, genera, species, subspecies, varieties, or landraces. Genetic diversity can be measured by different methods, ranging from the more traditional phenotypic approaches—in which the plant’s physical appearance is assessed through field or greenhouse observations—to modern methods that assess genetic diversity at the molecular level.

Crops are thought to harbor less genetic diversity than their wild relatives, since selection by humans during domestication causes a loss in genetic diversity—a process also known as a domestication bottleneck. Modern breeding efforts often lead to a further decrease in genetic diversity because of very strong selection pressures and the development of phenotypically uniform lines or varieties for commercial production. However, crop-breeding programs rely heavily on genetic diversity among different cultivars and crop wild relatives as a “genetic resource” for improvement. Also, the cultivation of a diverse array of varieties can be a promising preemptive strategy for disease management. Enhancing crop genetic diversity in agricultural production systems through the cultivation of a wide range of crops and varieties is increasingly recognized as a sustainable way to achieve food security, especially in areas with frequently fluctuating environmental conditions (see photo). However, as agriculture in most countries is being increasingly intensified using relatively few high-yielding and uniform varieties, crop genetic diversity is in decline, and much has already been lost. Increase of this decline in genetic diversity has been termed genetic erosion and is also caused by several other factors, including environmental and socio-economic changes.

Genetic resource conservation programs exist that aim to counteract the loss of crop genetic diversity through in situ or ex situ approaches. In situ efforts attempt to protect habitats of crop wild relatives or encourage farmers to grow and evaluate a great diversity of crops and varieties. Ex situ conservation efforts consist of national gene banks or botanical gardens, as well as international collections such as the gene banks of the Consultative Group on International Agricultural Research (CGIAR) or international backup collections such as the Svalbard Global Seed Vault.

Since crop genetic diversity is geographically unequally distributed around the globe, there is a considerable interdependence among nations and across generations for these genetic resources. As a result, the International Treaty on Plant Genetic Resources for Food and Agriculture (ITPGRFA) has been set up to govern international exchange of genetic diversity of many major crops, as well as ensure fair and equitable benefit sharing. Since humans exert much of the selection pressures that determine the amount of genetic diversity present within a crop’s gene pool and genome, traditional knowledge of crop diversity and production systems is an important source of information and
CROP ROTATION

Crop rotation involves the planting of different kinds of crops, in a time and space sequence, given the differential attributes of the crops, whereby a number of agronomic benefits are obtained. Although crop rotation is an age-old practice, it generally went out of practice in the late 20th century with the widespread introduction of agrochemicals and advanced technologies in agriculture, although some observers advocate a return to crop rotation on ecological grounds. In most countries, including those of the developing world, crop rotation had been a response to the growing population pressure on land resources, allowing lesser and lesser room for leaving the land fallow for a set number of years. Fallowing (leaving cultivated land unseeded and unplowed for one or more growing seasons), practiced since the earliest years of crop cultivation, allowed the soil to recuperate and also break the life cycle of crop pests and diseases. Allowing vegetation to grow on the fallow land allows for the accumulation of organic matter and other kinds of nutrients for the next cropping.

Crop rotation is a complex operation. It has a cultural context and is specific to agricultural systems. It should also be tailored to the social, economic, and environmental attributes of a place. Sorting of the crops grown in a particular area for the purpose of rotation also poses a challenge. Crops should be grouped, for instance, as grass crops or leguminous crops; those with shallow or deep roots; those with dense or light foliage; those with fibrous or large roots; those with allelopathic or nonallelopathic properties; those that need lesser or greater amounts of light, water, nutrients, and space; and those with susceptibility to the same pest or disease.

As chemical fertilizers, herbicides (weed killers), insecticides (pest exterminators), and treatments for plant diseases became more readily available, it became more attractive to use such convenient technologies in place of going through the drudgeries of rotating crops. The honeymoon with high-input farming did not last long. The skyrocketing prices of agricultural inputs (e.g., chemical fertilizers, herbicides, and insecticides) for the past 20-plus years have put crop farmers in a difficult situation. The resistance of some weeds, diseases, and pests to chemicals has posed a serious challenge to their effectiveness. Moreover, the environmental effects of the use of agrochemicals have also been reaching disastrous proportions. Although many farmers have continued to use agrochemicals, unimpressed by the looming environmental threats, there is an urgent need to revert to old and environmentally friendly farming practices, such as organic farming (which relies on crop rotation) and integrated pest management. A complete or partial shift to crop rotation could ensure agricultural sustainability as well as environmental sustainability.

The technological challenges that discourage the return to crop rotation are not limited to the relative ease in the use of artificial inputs. Most farming technologies are fit for monocropping and the application of chemical inputs. A return to crop rotation may involve redesigning of farm machines and implements. Even more intractable in this regard are the government policies and institutional arrangements, which are geared to high-input agriculture.

as such needs to be considered in efforts that aim to conserve crop genetic diversity.

Hannes Dempewolf

See also Agricultural Intensification; Agrobiodiversity; Agrofoods; Biodiversity; Biotechnology and Ecological Risk; Centers of Domestication; Conservation; Cultural Ecology; Domestication of Plants; Sustainable Agriculture

Further Readings


Crop rotation has several benefits that are provided by the varying botanical characteristics of the crops. It can improve the physical conditions, organic matter, and plant nutrient content of the soil; maintain the balance in the distribution of nutrients in the soil; enable nonchemical control of weeds, diseases, and insect pests, thus reducing reliance on pesticides; facilitate soil moisture management; make low-cost animal feed available; and enhance crop diversification (Figure 1).

For the best results, grass crops should not be followed by grass crops. When legumes are included in the rotation, their nitrogen-fixing and other characteristics will replenish the soil nutrient content depleted by the continued growing of grass crops. Since different legumes have different nitrogen-fixing capacities, crop selection should be an important consideration. Soil nutrient augmentation benefit also comes from other rotation crop properties. Deep-rooted plants could tap leached nutrients from the lower soil layers and return them to the topsoil after they die. The green manure from cover crops plays a similar role. Deep-rooted crops also improve the structure of the soil, increasing the soil’s aeration and water-holding capacity. Besides the advantage cover crops have in providing lower-cost feed for animals, they help improve the moisture-holding capacity and enhance resilience to drought. Moreover, by protecting the soil from being splashed by rain droplets and from being taken away by runoff, cover crops alleviate the problem of soil degradation.

Pests typically have specific host crops. When rotation is planned for the purpose of pest control, it is those crops that are not susceptible to the particular pest that should be planted. The crops can be of any type as long as they help break the life cycle of the pest, starve and kill it, or drive it away. The benefit of crop rotation for
weed control comes at least from three crop properties: (1) crops with dense foliage deny the weed sunlight and kill it; (2) some crops with allelopathic properties produce chemical substances that may inhibit the growth of weeds; (3) and those crops that could take more light, nutrients, and space could beat the weeds and eradicate them.

Crop rotation entails a significant change in the cropping pattern. The timetable and the types of crops grown may not be well-adjusted with the existing systems of production, marketing, and consumption. An agribusiness specializing in a particular crop and supplying a preplanned quantity of its produce to the market could face problems of irregularities in prices or supply of inputs. When this involves global markets, the problem becomes more complex. As crop rotation implies some degree of diversification, agribusiness should look for new markets for the alternative crops grown. A return to crop rotation does not cause as much inconvenience for subsistence farming livelihoods as it does for commercial agriculture. For either system, larger plots could
be divided into two or four segments and different types of crops grown in each plot (see photo). This can be rotated seasonally or annually depending, among other things, on the extent of nutrient depletion and the length of time needed to break the life cycles of pests.

Yohannes Aberra

See also Agriculture, Preindustrial; Crop Genetic Diversity; Organic Agriculture; Pest Management; Shifting Cultivation; Sustainable Agriculture

Further Readings


Cross-border cooperation (CBC) is a response to the challenges that international borders pose to the surrounding areas. It is the process in which neighboring local and regional actors engage in multifaceted cooperation across state borders to find mutually beneficial solutions to common problems that cannot be adequately addressed in a national framework alone. The primary goal is to transcend the barrier function of state borders.

CBC practices first appeared in Europe in the 1960s, and by the 1990s, they had become an integral part of the European Union integration process. Globalization flows have engendered the emergence of such practices in other parts of the world as well, most notably in North America and East Asia. Economically, CBC has the potential to enhance the development of the borderlands by allowing economic actors to take advantage of opportunities situated on both sides of the border. Culturally, CBC can break down negative stereotypes by promoting good neighborly relations. Politically, CBC could enhance the democratic process by bringing decision making closer to the borderland inhabitants. In achieving such objectives, CBC faces the considerable task of developing integrative mindsets among a variety of local actors to allow them to identify shared interests necessary to create common spaces of living.

CBC departs from traditional intergovernmental cooperation, allowing subnational authorities and civil society actors to engage in direct interaction across state borders. Nonetheless, these practices do not entirely bypass the national governments. They tend to be project specific rather than forming a comprehensive territorial strategy for the management of social life in the borderlands. CBC is generally oriented toward building cross-border institutions that can provide enabling frameworks for cooperative actions. The institutionalization of CBC typically assumes the form of multilevel governance networks that can involve local, regional, and national governments, supranational institutions, development agencies, universities, chambers of commerce, and nongovernmental organizations, all interacting in a loosely coupled relationship based on coordination and negotiation rather than top-down subordination.

Cross-border regions, commonly known in Europe as Euroregions, are currently the most common and complex spatial frameworks for CBC. They are territorially delineated regions that straddle state borders, where cooperation can be organized irrespective of state borders, to the benefit of the civil society (Figure 1). They can have governing institutions, such as councils and secretariats, and symbols, such as logos and flags. In practice, the persistence of national particularism among many actors involved in CBC has prevented cross-border regions from becoming meaningfully integrated spaces of living.

CBC processes aimed at cross-border spatial integration are often in a state of tension with the nation-state’s sovereignty demands involving a border-containment territorial logic. These conflicting territorial logics often prevent CBC from reaching its full potential. Nonetheless, the overall impact of CBC has been to unsettle the long-established meaning of state borders as ultimate
Cultural ecology is a subfield of geography that is shared with anthropology and other allied fields. It positions itself in the center of the human-environment or “nature and society” tradition in geography (which itself is considered to be the core and uniqueness of geography) and concerns itself with the way in which humans are integral components of the environment and “ecology” of a particular place. It has long focused on the historical context of human-environmental relations, both the human and the natural history of place and how humans have adapted to nature, as well as how humans have modified nature for their lines of defense and to accommodate increased contact and gateway functions.

Gabriel Popescu

See also Borderlands; Borders and Boundaries; Governance; Regional Governance; Sovereignty; Transnationalism

Further Readings


Figure 1  Euroregions situated between Romania, Ukraine, and Moldova

Source: Map created by Cristina Scarlat, Center for Advanced Spatial Technologies at the University of Arkansas.
Cultural ecology focuses on process rather than material culture. Key elements of the sub-field are a focus on marginalized and/or vulnerable peoples (both in the developing periphery and more recently in the developed core), a concern with the impacts of economic development on local livelihoods, a focus on threatened landscapes as well as (agro)biological conservation, an interest in local knowledge, and a focus on rural and agrarian systems, although recently, urban space has received significant attention.

Cultural ecology’s research approach has always focused on intensive empirical fieldwork, often long-term and place based. Research questions are field based and use theory to inform, but they are not theory driven. Though originally focused on deep ethnographic fieldwork, often combined with historical analysis, cultural ecology has consistently used mixed methods, ranging from an incorporation of quantitative analysis (e.g., survey research, including approaches from microeconomics) to participatory methods (e.g., mapping, transect walks, seasonal calendars) and research methods from the physical sciences (e.g., soil surveys, plant measurements and surveys).

More recently, cultural ecology has increasingly used discourse analysis and other tools from critical theory, as well as geospatial technologies (e.g., the analysis of remotely sensed images, use of global positioning systems [GPS] as a tracking tool, and use of geographic information systems to incorporate information from different sources and scales), in its methods suite. The use of the latter technologies has generally reduced field time for practitioners, but it has opened up the ability to scale up from what had been a focus on local-level studies.

Historical Development

Cultural ecology emerged as a subfield of geography in the 1960s and reached its peak in the 1970s and 1980s. An Association of American Geographers (AAG) specialty group, the Cultural Ecology Specialty Group, was formed in 1980 to represent and build cohesion in the sub-field, which at that time focused on topics ranging from prehistory to Third World development and from environmental to economic issues. The establishment of the new specialty group recognized the field’s interdisciplinarity and eclectic nature.

Cultural ecology, having originated on the basis of the ideas of the cultural anthropologist Julian Steward, as presented in 1955, has long been interested in adaptation—not just how humans adapt to their environments but how they have modified their environments to suit their needs. It has been quite concerned with human agency and with the human capacity to manipulate the environment, and it interprets adaptation as choice and not as environmentally determined. William Denevan’s work is seminal in this regard. Practitioners have worked hard to counter the ideas of environmental determinism, which were so destructive to the discipline of geography.

The rise of the quantitative revolution in geography was a great challenge to all subfields that did not embrace it, but especially cultural ecology, which was focused on localized case studies and was not interested in generalizing social science theory. The subfield was seen as marginal to the discipline during this era, although it remained active and provided a continual body of knowledge on changing livelihoods throughout the then Third World. It embraced systems ecology in the 1960s and 1970s, and some of its key research (e.g., Bernard Nietschmann’s work) used energy flows to explain human livelihoods.

Although in its early days, cultural ecology was not explicitly political, it was never apolitical, although there was a persistent myth that it was. As Paul Robbins has noted, where cultural ecology did fall short was perhaps in its tool chest to explain the larger processes in which marginalized peoples of the Third World found themselves. Globalization and neoliberal structural reforms were accelerating change around the world, and the use of various approaches from critical theory permitted the unpacking of these structural layers.

The Rise of Political Ecology

Cultural ecology as an identifiable subfield fell out of vogue during the 1990s, in part because of the critique regarding the lack of political contextualization and in part because of the rise of
CULTURAL ECOLOGY

postmodern and poststructuralist tools to explain the larger processes of change. Political ecology, a cognate of cultural ecology that wholeheartedly embraces social theory, rose to prominence in the 1990s as the dominant approach in people-environment geography. Cultural ecology has long been driven by questions in the field and not by theory, whereas political ecology in general is driven more by theory, hence political ecology came to dominate at a time when theory was seen as the way to explain process.

During the 1990s, although cultural ecologists remained active, political ecology dominated, and many long-time cultural ecologists felt marginalized and turned their research in other directions. Some turned back to their earlier roots in physical geography; others embraced what began as historical ecology. Although some observers believe that cultural ecology has been subsumed under political ecology, there is a vibrant and growing community of young, neocultural or contextualized cultural ecologists practicing in geography today, many of whom eschew labeling their work as either cultural ecology or political ecology. These researchers have often trained with noted scholars from either classic cultural ecology or newer political ecology and are now setting forth on their own pathways and, in the process, are redefining the field. In reality, cultural ecology and political ecology are a continuum that as a whole is so rich because it is inclusive and can accommodate a range of approaches.

In recognition of the rise of political ecology, the AAG specialty group was renamed the Cultural and Political Ecology Specialty Group in 2002. While the name change debate left some traditionalists marginalized, it engaged many younger scholars, and the specialty group remains one of the largest and most active in the AAG, with more than 600 members in 2008.

Cultural Ecology Today

The subfield of cultural ecology continues today as the core of a vibrant field within human-environment geography, although practitioners may not call their work cultural ecology (indeed, many have long resisted the label), even in the heyday of cultural ecology. Much of 21st-century cultural ecology is labeled as livelihood, development, or landscape studies, all of which are integrative and cross-disciplinary fields, as are geo-archaeology, environmental history, or historical ecology, three other identities cultural ecologists have embraced. What continues to set apart this neocultural ecology from political ecology is a continued focus on local food and subsistence systems, and especially the environmental knowledge systems that support these. Still others include a much greater focus on the ecological and work in the field they call human ecology.

The newest turn for cultural ecologists is that they are working closely in or with the “land change” or “sustainability science” (including the LUCC, or Land Use and Cover Change) community, where the bigger questions of global environmental change are analyzed using a suite of methods, including those that cultural ecologists have long been using. Local, field-based cultural ecology often informs larger-scale models. Cultural ecologists, given their long-term interest in adaptation, have also become involved with climate change investigations and are positioned to contribute significantly to the field. Along with these, cultural ecologists have embraced looking at scale, recognizing the limitations of the local and the need to find ways to scale up from the local to the global.

Cultural ecologists, given their wide range of interests and contributions, publish their findings in a wide range of journals, including the key geographical ones (especially The Geographical Review) and some anthropological ones (e.g., Human Organization), but perhaps the greatest concentration of this type of research is found in the interdisciplinary journal Human Ecology. For a detailed review, the reader can consult the two surveys of the subfield written for the two benchmarks of the discipline of geography: Karl Butzer’s chapter in Geography in America and Tom Bassett and Karl Zimmerer’s chapter in Geography in America at the Dawn of the 21st Century. These two in-depth publications list most of the cultural ecological work completed during the past several decades. More recently, Zimmerer has written a series of articles published in Progress in Human Geography, in which he places current cultural ecological research within the dominant and dynamic strands of research to
which the subfield is contributing substantially thus far in the 21st century.

Antoinette M. G. A. WinklerPrins

See also Class, Nature and; Critical Human Geography; Critical Studies of Nature; Environmental History; Ethnicity and Nature; Gender and Nature; Human Ecology; Land Use and Cover Change (LUCC); Nature; Nature-Society Theory; Political Ecology; Social Construction of Nature

Further Readings


What Is Culture?

One of the most vexing aspects of cultural geography is the difficulty of defining its object of study: culture. In this, cultural geographers and cultural anthropologists share a similar dilemma, with the result that various attempts to define culture have long crossed disciplinary boundaries between geography and anthropology. In fact, it was the anthropologist Clifford Geertz (1973), drawing on earlier work by the sociologist Max Weber, who famously asserted, “believing . . . that man is an animal suspended in webs of significance he himself has spun, I take culture to be those webs” (p. 5). Culture has been defined as a whole way of life, a manner of thinking and doing things that encompasses an entire group of people. A more restricted definition of culture is synonymous with the tangible things that particular groups of people produce: material culture. Culture can be used in a way that indicates high culture, a society’s artistic and intellectual output. Finally, culture can have a temporal meaning, as opposed to being a static phenomenon, when the term is used to indicate the progressive advancement of a group.

In several of these definitions, culture is a thing. According to some understandings, culture is an attribute or a set of distinct qualities possessed by a group of people; in others, culture is literally things, the distinct material, artistic, spiritual, and intellectual output of a particular society. Definitions of culture as a thing have at least three problematic aspects associated with them. First is their tendency to accrete just about
anything. Even the design, technology, and location of the proverbial kitchen sink can be thought of as cultural. There is a loss in analytical precision when the definition of a term can potentially include everything. Second, if culture is a distinctive attribute of specific human societies, it is tempting to view the world as a patterning of diverse cultures that are juxtaposed side by side, as with a jigsaw puzzle. Such definitions cannot account for the persistent mobility, intermixing, and fluidity of culture. The third problem with approaches to culture as a thing is that culture is relegated to an intermediary role. Cultural output merely provides clues to deeper patterns: social, economic, political, and so on. Culture is thus epiphenomenal, rather than a powerful force in its own right.

In other definitions, culture is approached as a process. Culture is understood to involve ways of communicating, expressing, making meaning, and representing external reality to ourselves and others. Some scholars view cultures as language-like in their interactive and symbolic structure. Others have a less structured view of culture as a process, emphasizing instead culture’s unstable fluidity. These approaches have been critiqued for at least two reasons. First is their tendency toward “realm-ness,” for lack of a better term. In these approaches, culture is viewed as an arena in which activity occurs, distinct and separate from other arenas, such as the political, economic, environmental, and so on. Thus, culture is artificially cordoned off and reified as a separate realm, when, in fact, there are cultural aspects to many activities not typically seen as having a cultural dimension. Indeed, it is in the apparently obvious, but misguided, notion that some activities are acultural that culture can work in extremely powerful ways. To assume that politics or economics exist apart from culture, even in the abstract, is naive. Second, approaches to culture as a process often provide too vague answers to the straightforward question “What is culture?” Not only does this make research difficult, it can render practices deemed cultural difficult to hold accountable.

Frustration with the difficulties associated with culture has led some contemporary cultural geographers to assert that we inhabit a postcultural era, that there is no such thing as culture, or that culture is such a plastic idea as to be dangerous. Others counter that while many definitions of culture are possible, it is vital to operationally define culture so that the important work of cultural analysis may continue.

Major Divisions Within Cultural Geography, by Place

Though we speak of cultural geography as if it were a coherent, unified subfield, it is not. This arises partly from the different conceptualizations of culture discussed above. Another major division in cultural geography concerns the site in which the distinct cultural geography traditions arose. The principal place-based division within cultural geography is between American and European approaches. It is important to state early on, however, that these two traditions in Anglophone cultural geography themselves are the product of much trans-Atlantic intermixing. As with all things cultural, it is important not to underestimate the importance of intellectual cross-place fertilization. Cultural geography is no exception.

U.S. Versus British Traditions

In the United States, cultural geography is known both for its carefully descriptive nature and its long history of discomfort with this aspect of its approach. The work of the best-known American cultural geographer, Carl Ortwin Sauer, distills this tension. Sauer’s approach to cultural geography centralized the landscape as the product of the action of human society on nature, viewed over time. Human culture literally sculpts the natural world into its visible contours. The task of the cultural geographer, then, is to meticulously observe and record the features of the landscape and from these to deduce the features of the societies that shaped it. This is, however, no subjective interpretive task, though Sauer retained a respect for the aesthetic dimensions of landscape interpretation. Troubled by what he understood as a lack of scientific rigor on the part of earlier cultural geographers, Sauer insisted that a systematic methodology be employed to record, classify, and link the impress of human culture on
the natural environment. He called this methodology *morphology*. Sauer’s many students, and his approach more generally, have influenced generations of American cultural geographers, whose work is characterized by its rich description of material culture, the central role of field-based research, and a deep concern with scientific rigor in data collection and analysis.

Institutionalized European cultural geography predates American cultural geography; indeed, early American cultural geographers were trained by European scholars and drew on their work. Sauer, for instance, incorporated the regional descriptive tradition of the French geographer Paul Vidal de la Blache as well as the German political geographer Friedrich Ratzel’s emphasis on the separate yet interrelated natural and human worlds. It is British cultural geography, however, that has developed the highest profile among the European cultural geography traditions. British cultural geography emerged from a concern with the shaping effect of broad transformations in society. Thus, the work of important early British cultural geographers, such as W. G. Hoskins, was referred to as *landscape analysis* rather than cultural geography. Hoskins sought to understand the English landscape as a text of sorts, from which could be read the transformation from a rural peasant society to a modern, urban, industrial society. The driving force in Hoskins’s understanding was the march of social history, including changes in transportation, industry, and agricultural technology, which used as well as shaped the landscape. Compared with its American counterpart, British cultural geography has less encyclopedic description and emphasis on scientific methodology and relatively more emphasis on social theory. Today, as well as in the past, however, these distinctions are oversimplified, and a vibrant trans-Atlantic exchange of ideas is ongoing.

**Non-Anglophone Traditions**

The oldest cultural geographic traditions are found in ancient Greek and Roman, as well as Muslim and Chinese, scholarship. Because the inhabitants of the ancient world were far more place bound than most people are today, speculation about distant lands and their inhabitants was common, while empirical accounts were rare. Explorers constituted the earliest de facto cultural geographers, and they disseminated their findings to a wider audience. It is hypothesized that cultural geography as well as cultural anthropology are important dimensions of imperial societies. Thus, it is no accident that Chinese, Muslim, Greek, and Roman civilizations—all of them expansionist empires of the ancient world—counted what would now be classified as cultural geographers among their most visible intellectual ranks.

Today, for better or for worse (and as with much social science), Anglophone scholarship dominates the global marketplace in cultural geography. To complicate matters, disciplinary divisions can be very different from place to place. Scholars who might be classified as cultural geographers in the Anglophone tradition can be found in departments of anthropology, urban sociology, and planning, and even in government offices such as census bureaus, in non-Anglophone contexts. Finally, some academic traditions place a strong emphasis on activism, something that is fairly uncommon in Anglophone academia.

**Major Divisions Within Cultural Geography, by Approach**

Looking at different place-based approaches to cultural geography is one way to understand the differences within the subdivision and how they came about. Another way to frame these differences is by intellectual emphasis. Using this lens, at least three different approaches can be isolated: empiricist, humanistic, and critical. Three caveats are in order before proceeding. The first is that these labels are merely a convenient short-hand to convey more or less significant intellectual differences. Second, in reality, there is much overlap and mutual conversation among these three approaches. Third, the divisions discussed earlier concerning diverse definitions of culture and place-based distinctions in cultural geography traditions thread into these intellectual distinctions here, such that a tightly braided skein results. In practice, it is difficult (and rather boring) to attempt to isolate the multiple ways in
which cultural geographic traditions are different from one another; rather, the more interesting task lies in identifying constellations of productive focus within the larger subfield and working from those.

**Empiricist Approaches**

The foundational cultural geographers in Anglophone cultural geography were largely concerned with material culture. Empiricist cultural geographers particularly emphasize the visual features of the landscape, those aspects that are apparent to the human eye. Empiricist cultural geographers address questions such as “What traces have societies past and present left on the landscape?” and “What do these traces tell us about those societies: their political, economic, and social structures, the transformations in these structures, and the ways these societies have adapted to and modified the natural world?”

The type of housing built by a society, for instance, provides clues about that culture that one can decipher if one is attentive enough to the details. What building materials are used? How is the roof pitched? How large is the typical dwelling? What shape is it: round, square, triangular, or rectangular? Is human space shared with animals, or are living spaces for human and animals carefully separated? Are there common spaces among family dwellings, or are spaces entirely privatized and spatially marked as such? Are dwellings clustered together, isolated, or arranged along roadways? By definition, empirically oriented cultural geographers centralize the importance of long-term immersion in the field as vital to recording and understanding the level of detail required for accurate analysis.

The distinction between folk and popular culture has long been a mainstay of this approach. The empiricist strand of cultural geography was solidified during an era marked by the fascination with understanding the spatial dimensions of the transformation from rural peasant society to urban industrial society. Folk culture denotes the traditional practices, beliefs, and material cultural elements of rural people. Popular culture, in contrast, speaks to modern, urban, and nonlocal practices, beliefs, and material cultural elements.

Though it is hardly possible today to find purely local folk cultures untouched by modern influences, cultural geographers in this genre tend to retain some form of distinction between folk and popular cultures, though it is usually not delimited as starkly as it once was.

The work of empiricist cultural geographers has been likened to that of a detective decoding pieces of a larger puzzle. Thus, though all three of these approaches can be said to view the landscape as a sort of text, in the empiricist cultural geography tradition, the landscape is a book that, if read carefully, will yield up a straightforward story about the culture that wrote it.

**Humanistic Approaches**

*Humanism* can be simply defined as a strong belief in the redeeming quality of the human spirit. Humanistic scholars focus on the needs, abilities, and freedoms of human beings. Though most commonly found in the arts, art history, literature, and philosophy, some cultural geographers also draw from a humanistic tradition. Their work primarily examines human connections to place. The notion of dwelling, or establishing a deeply rooted connection to place over time through repeated activity, is important to humanistic cultural geographers. The affective dimensions of this connection to place can be summed up in the cultural geographer Yi-Fu Tuan’s notion of *topophilia*: literally, love of place. Poetic, artistic, literary, and spiritual ways of establishing and representing landscape, belonging, home, sacredness, place, lifeworld, and states of being such as childhood are found in humanistic cultural geography.

**Critical Approaches**

Cultural geography is not immune to broader developments in social theory; indeed, cultural geographers have contributed to these developments, particularly through an increasing prominence of space and place in social theory more generally. In part as a reaction against the lack of attention to social theory by empiricist cultural geographers, as well as what was seen as the overly quantitative tone of mid-20th-century geography in general, critical cultural geographers...
Critical cultural geographers are noted for their attention to the spatial expression, representation, and negotiation of broader power relations in society. The hidden aspects, silences, and omissions are every bit as important as those aspects of the cultural world that are manifest. In addition, a close attention to social justice—through exposure of inequality and the activism directed at rectifying it—is a hallmark of this approach.

Take, for instance, the representation of an idyllic landscape constituted of rolling hills, a rustic cottage, and peaceful cows. A critical cultural geographer would focus on how and why the landscape was produced in this fashion. Where are the laborers who planted, pruned, and mowed the vegetation? What of the woman who swept the path and tended the garden? Why are such landscapes, devoid of people and containing certain “natural” visual elements arranged just so, considered so appealing? Addressing such questions, and proposing solutions to the spatial injustices that produce the visible and invisible

Jerusalem’s Wailing Wall is a sacred site where Jews from all over the world come to pray. Cultural geographers find such sites to be incredibly rich. The Wailing Wall constitutes an axis mundi, or anchoring point, around which the entire city is referenced for Jerusalem’s Jewish residents. A mixture of popular and folk cultures is evident in the styles of dress. Women and men pray at separate sections of the wall; the women’s section is much smaller and more crowded than the men’s section. Even the spatiality of the body is affected by the site’s sacredness. The faithful refuse to turn their backs to the wall, as doing so would be disrespectful, and so they leave the site by walking backward.

Source: Author.
worlds around us, is at the heart of critical cultural geography.

Cultural Geography Today

Cultural geography is today one of the most dynamic subfields of human geography. Many students find cultural geography’s breadth of topics exciting. In addition, the sorts of questions that can be addressed using a cultural geography approach are among the most crucial facing us today: Globalization, identity, mobility, and human-environment relations have all been examined by contemporary cultural geographers.

Cultural geography’s traditional themes are still present but share the stage with new approaches. Among these are nonrepresentational ways of apprehending cultural phenomena, which attempt to circumvent the contemplation, interpretation, and representation of reality. Rather, nonrepresentational theory emphasizes performative, embodied practices as valid forms of knowledge. Nonrepresentational cultural geography has focused on movement, such as dance, play, and exercise, as creating an experiential landscape that constructs and deconstructs as the subject moves through space. Rather than the stability provided by representational strategies, nonrepresentational cultural geography emphasizes the inherent instability of lived spatiality.

Another area of development in contemporary cultural geography involves emotions. Cultural geographers in this tradition examine the role of fear, hate, love, and other emotions in how we experience, represent, and contest spatiality. Related to this field are those cultural geographers who explore sensual geographies beyond the visual. Their incorporation of smell, taste, touch, and hearing broadens cultural geography’s long emphasis on solely the visual aspects of spatiality.

Contemporary cultural geographers also include different subjects of study. Cultural geographies of the elderly, of children and youth, racialized minorities, and disabled people have become increasingly important. Geographies of pregnancy, prostitution, and queerness all focus on sexualized aspects of spatiality. Nature and animals, and the blurry boundaries between traditional human subjects of cultural geography, constitute another emerging area of focus for contemporary cultural geographers.

Patricia L. Price

See also Architecture and Geography; Art and Geography; Berkeley School; Body, Geography of; Children, Geography of; Class, Geography and; Consumption, Geographies of; Cosgrove, Denis; Critical Human Geography; Cultural Landscape; Cultural Turn; Difference, Geographies of; Discourse and Geography; Emotions, Geography and; Ethics, Geography and; Ethnicity; Ethnocentrism; Eurocentrism; Everyday Life, Geography and; Existentialism and Geography; Feminist Geographies; Fieldwork in Human Geography; Folk Culture and Geography; Gays and Lesbians, Geography and/or; Gender and Geography; Geographical Imagination; Humanistic Geography; Hybrid Geographies; Identity, Geography and; Indigeneity; Jackson, John Brinckerhoff; Landscape Architecture; Landscape Interpretation; Languages, Geography of; Masculinities and Geography; Media and Geography; Music and Sound, Geography and; Nationalism; Nonrepresentational Theory; Nonvisual Geographies; Orientalism; Phenomenology; Place Names; Popular Culture, Geography and; Population Geography; Postcolonialism; Poststructuralism; Queer Theory; Race and Racism; Religion, Geography and; Representations of Space; Sauer, Carl; Sense of Place; Social Geography; Sports, Geography of; Subaltern Studies; Symbolism and Place; Text/Textuality; Topophilia; Tourism; Tuan, Yi-Fu; Urban Geography; Vidal de la Blache, Paul; Vision and Geography; Whittlesey, Derwent; Writing; Zelinsky, Wilbur

Further Readings


The cultural landscape is a concept associated with human modification of Earth’s surface. Cultural landscape studies have been a prominent component of cultural geography over the past century. Despite a long-standing interest in the topic, geographers over time have debated the concept and associated theories and methodologies used to understand cultural landscapes. The term cultural landscape is associated with the depiction of the Earth’s surface in visual art, material artifacts such as buildings, and the area of territory itself. Furthermore, it is a term that is not exclusive to geography, as sociologists, anthropologists, and historians have used cultural landscape as well.

Early Development of Cultural Landscapes

The term landscape developed primarily from two words that display similarities and differences with one another. The Old English word landscaef referred to a clearing of forested land by humans or, to be more precise, the occupation or control of the “wilderness.” This term provides geographers with the basis for studying the human-modified environment, as well as the interaction of humans with the land. The second term, landschap, comes from the Dutch and is associated with the appearance of the land. This latter term was primarily linked to visual art, especially paintings. This allows for a certain “reading” of the landscape.

In geography, the concept of the cultural landscape has antecedents in the research and philosophies of French and German geographers of the late 19th and early 20th centuries. The German geographer Otto Schlüter, who is often credited with coining the term cultural landscape, held the landscape itself as the centerpiece for all geographers. The various forms and spatial arrangements create a bridge between physical geography and human geography, as the landscape itself is an object that both sides of the geographic dichotomy focus much attention on.

However, Schlüter’s approach to the study of landscape geography, or more precisely landscape morphology, focused only on the visible landscape and therefore did not dwell entirely on the meanings and symbolism behind the landscape. In other words, Earth should be studied from a bird’s eye view, where the roads, buildings, gardens, oceans, and forests create a visible image that is the focal point for geographical study. Others, however, felt that this limited focus on the visible landscape detracted from the pursuit of deeper meanings that could be discovered in the cultural landscape.

It is also worth noting that the German school of geography is credited with developing yet a third meaning for landscape associated with the German word landschaft. This term deals primarily with forms of landscape in a particular area. To the Germans, landscape studies were synonymous with regional studies.

Another aspect of the cultural landscape was the marriage of social and physical characteristics. The French geographer Paul Vidal de la Blache held that the natural landscape and the human landscape should be viewed as inseparable. To Vidal, every cultural group adapts to its natural environment in its own way and creates a unique region. A result of this adaptation is the formation of the cultural landscape, which is a reaction to the natural landscape. This process amalgamates two seemingly opposite ideas—society and nature—into one observable phenomenon, one that Vidal believed should be the focal point of geographic study.
Early research involving the cultural landscape in the United States goes back to the work of Carl Sauer and the Berkeley School of cultural geography. Sauer saw in the cultural landscape an expression of the material features of cultural groups. His famous paper titled “The Morphology of Landscape,” published in 1925, attempted to solidify the landscape as the fundamental concept of geography (Figure 1). Through this idea, cultures shape the natural environment to create the cultural landscape. Sauer’s idea brings together Schlüter’s and Vidal’s concepts—the former’s attempt to bring together the physical and the human and the latter’s focus on regional uniqueness. As Sauer (1925) stated, “The cultural landscape [is] fashioned from the natural landscape by a culture group. Culture is the agent, the natural area is the medium, the cultural landscape is the result” (p. 46). This view lent itself to the German term landschaft, as it stressed the importance of regional uniqueness.

Sauerian notions of landscape were seen as a reaction to environmental determinism, a theory that dominated early-20th-century U.S. human geography and that highlighted the ways in which the environment influenced humans. Sauer’s approaches to landscape focused on the interactions between humans and the environment,
providing emphasis on the role of human agency. Sauer did not ignore the impacts of the environment in the creation of the cultural landscape, which he emphasized was humanly produced. Despite Sauer’s attempts to integrate cultural traits, this view still held the cultural landscape as simply a material manifestation of the cultural group, without understanding the deeper meanings behind the human creations.

While some early researchers did examine the cultural landscape with a deeper focus, many of the late-19th-century and early-20th-century geographers did not dig deeper than the surfaces of the landscape. The cultural landscape was viewed as static and represented a politically neutral vantage point. Today, however, cultural geographers do not hold landscape as a politically neutral manifestation of society. Since the end of World War II, geographers have recognized that while humans create and manipulate the landscape, the landscape also influences humans. Aspects of the cultural landscape can give rise to a certain meaning to a particular cultural group, or more simply a “sense of place.” This sense of place in terms of landscape can create an experience that is shared by those whose daily lives are somewhat connected to the cultural landscape. By creating this shared experience, a centripetal force is exerted that can be used to unite a cultural group together and potentially stir up nationalist sentiment.

Early geographers took the cultural landscape at face value, examining the diffusion and location of items such as housing types. This type of research did not observe why there was diffusion but tended to simply categorize the cultural landscape; it has since been replaced by various tactics that seek to bring about a deeper understanding of the landscape.

Deeper Meanings in the Cultural Landscape

Many cultural geographers in the post–World War II era were heavily influenced by the work associated with the Berkeley School. For example, Wilbur Zelinsky and Fred Kniffen were among the many geographers who examined not only the occurrence of similar town designs but also their diffusion from a cultural hearth to other parts of a region.

J. B. Jackson is also prominently associated with studies involving the cultural landscape. He defined landscape as the portion of the Earth that can be comprehended at a glance. However, although he was trained in the visual and literary arts, Jackson became interested in the landscape as something that is lived in rather than a simply visual phenomenon. Jackson focused on the vernacular landscape, which consisted of everyday material items such as strip malls, fast-food restaurants, motels, and groceries. In this case, landscape was a reflection of human society and displayed a deeper understanding of that society. For Jackson, landscape, like art, needed to be “read” in order to gauge its deeper meanings. He likened the landscape to a book, one that is “always open before us.” However, while that book was open, we must “learn to read it.” And because different people—both within and outside a particular society—will read the landscape differently, this reading is often in the eyes of the beholder. As Donald Meinig notes, 10 different people—both within and outside a particular society—will read the landscape differently, this reading is often in the eyes of the beholder. As Donald Meinig notes, 10 different people looking at the same landscape could interpret it in 10 different ways. For example, a Walmart in a small rural town could perceivably be read in a number of ways by both those living in the town and those who are just driving through the town. This notion speaks to the fluid nature of the cultural landscape.

Research Approaches to Cultural Landscapes

Contemporary geographers have used various approaches to the study of the cultural landscape. Some geographers took a behavioralist approach to the cultural landscape, including the use of mental maps based on daily experiences. Mental mapping worked well at the local scale, where individuals’ daily interactions with the cultural landscape produce a sense of place familiar to the particular individual. These mental representations create an understanding of spatial distribution that relies more on relative location than on absolute location. Another approach was analysis of the symbolism of a place in light of past historical experiences. For example, European colonizers were attracted to certain parts of the “New World” partly because those areas resembled
their homelands—an area and landscape with which they were familiar.

Another aspect that follows a humanist tradition is the concept of “reading” the landscape. In this view, the landscape is more than something to just look at; it is also a representation of the human interaction with nature. In other words, the landscape is a social construction of the world according to a particular cultural group. With this in mind, the cultural landscape not only is based on the individual’s perception but is a representation of society as a whole.

Geographers also added the element of time, placing emphasis on the evolution of a natural landscape into a cultural landscape. This notion includes the temporal changes of various cultures and the effects that each culture has on the landscape to form a palimpsest. For example, the introduction of British and Dutch culture, as well as South Asian culture, to the indigenous African culture had a profound impact on the cultural landscape of South Africa. Another example is the changing demographics of the central cities of the United States, as many cities have experienced a shift from being predominantly white to predominantly black. In addition, immigration has also greatly affected the landscapes of urban areas, as with the Cubans in Miami, Irish in Boston, or Japanese and Chinese in San Francisco.

Another perspective focuses on the symbolism behind the creation of the cultural landscape. Religion, language, race and ethnicity, gender, sexuality, and class all shape the cultural landscape. For example, Islamic mosques in Singapore, French place names in Louisiana, Chinatown in Vancouver, the Castro District of San Francisco, and the favelas of São Paolo have all been shaped by the ideologies and practices of specific cultural groups. Furthermore, these landscapes often mask conditions that promote stereotypes or obscure harsh conditions. In the view of Don Mitchell, the landscape can and often does “lie” to us, conveying one set of meanings as natural or inevitable (i.e., those of the dominant class) and hiding other, equally valid meanings generated by politically marginalized groups.

Geographers have attempted to understand the multiple, often contradictory, meanings embedded in the cultural landscape. This task involves exposing the power relations behind the construction of particular landscapes, including race or ethnicity, gender, sexuality, and class. For example, in the case of cultural landscapes that memorialize historical events, many geographers investigate the question of “whose” history is being depicted. These cultural landscapes typically display a history that promotes the dominant culture’s version of events. However, there are also examples of minority cultural groups that have developed landscapes that depict their own history and challenge the dominant cultural landscape, such as the dedication of street names in cities in the United States in honor of the civil rights activist Martin Luther King Jr.
In short, geographers have embedded landscapes within the complex political and social dynamics of societies, seeing them as both reflective and constitutive of the organization of society. In this sense, the study of cultural landscapes has become highly attuned to politics, justice, and inequality. For example, cultural landscapes that depict a historical event typically do so from a particular point of view—most often the view of the elites. Even the language of the term has been challenged, with *space* or *social space* being preferred over *landscape*. Furthermore, the use of the word *culture* is also seen as problematic. Marxist geographers see culture as heavily manipulated by elites as a way of maintaining their power. Therefore, cultural landscapes represent a particular elite culture. Landscapes often naturalize social relations, hiding the system that produces them. To study landscapes, therefore, it is necessary to unveil the specific power relations that go into their making, including not only the material distribution of tangible phenomena but also the ideology of landscape. In this light, there can be no “objective” reading of landscape, only different power-laden views that serve different social interests.

*Daniel McGowin*

*See also* Chorology; Cosgrove, Denis; Cultural Geography; Ethnicity; Hettner, Alfred; Jackson, John Brinckerhoff; Landscape Architecture; Landscape Design; Landscape Ecology; Landscape Interpretation; Landscape Quality Assessment; Landscape Restoration; Land Use Analysis; Meinig, Donald; Mitchell, Don; Palimpsest; Place Names; Sauer, Carl; Spaces of Representation/Representational Space; Symbolism and Place; Urban Land Use; Vidal de la Blache, Paul; Whittlesey, Derwent; Zelinsky, Wilbur

**Further Readings**


**Cultural Turn**

The term *cultural turn* refers to an intellectual recognition, within human geography and the wider human sciences, that all claims about the world and knowledge are mediated by culture. The logic of the cultural turn centers on the argument that no representation merely reproduces what it claims to represent. Representations construct, as much as they claim to explain, the person, place, or thing to which they refer. The implications of the argument that knowledge produces the world, rather than simply mirrors it, are profound. The argument entails that to make, and/or accept, any claim, we need to attend to two interrelated terrains. First, we need to attend to how knowledge claims are complex products of particular contexts and specific perspectives and are shaped by forces such as history, gender, class, race, and location. Second, claims about objects and the world affect both the possibility of identifying objects as such and the ways objects and the world are made meaningful. Never simply passive, knowledge emerges from, and brings into being, dynamics of difference, power, effect, meaning, and influence. Questions of language, politics, history, identity, and interpretation (i.e., culture) are thus central to the knowledge and constitution of “reality.”
It follows that careful scholarship must reflexively attend to the interpretive and complex constitutions of knowledge. Denying the influence of culture on what and how claims are made excludes how knowledge and meaning making are enmeshed within practices of power. Arguments made from the cultural turn entail the impossibility of objective knowledge, for claims to objectivity elide their contingency and partiality. Naturally, claims about the world can be made in the realms of science, ethics, and politics, but they must self-consciously recognize their embeddedness within multiple and situated positions. Imperatives of heightened reflexivity render knowledge making not only more complex and humble but also more attentive to its construction and effects. Debates surrounding the cultural turn have shaken, fundamentally, the academic humanities and social sciences for the past three decades. While the turn to culture and its epistemic implications has been enormously influential in shaping the possibilities of humanities and social science research, the repercussions of acknowledging the close relationships of language, power, and knowledge continue to be ardently felt and debated.

The turn to culture emerges out of a history of thought whose lineages can be traced through the influence of both the social writings of Karl Marx and the perspectivism of Friedrich Nietzsche. Significantly for the cultural turn, the work of Marx and Nietzsche was central, in the 1960s, to poststructuralism, which argued that language is necessary for thought. If language and text are the constitutive apparatuses by which the world can be known, there is nothing that can be said of the world that is outside of, or beyond, language and texts and, hence, culture. This view, and its effect on philosophy and social theory, became known as the linguistic turn. The cultural turn extends the insights of the linguistic turn to thinking about how cultural practices shape both our being in the world and the possibilities of responding within discourses of power, subjectivity, governance, and accumulation. From these possibilities, understood as opportunities, the interdisciplinarity of cultural studies emerged. For human geographers, the influence of cultural studies is most keenly felt in cultural geography, although the cultural turn has also transformed how subdisciplines such as economic, social, political, and environmental geographies are framed and engaged.

Often prefixed with the term new to denote the discourses emergent in the early 1990s, new cultural geographies, new economic geographies, new urban geographies, and so on work in a vast range of ways to reimagine traditional areas of research under the influence of new theoretical approaches, including postcolonialism; consumption; mobilities and their receptions; social constructions of identity with an attention to race, gender, and sexuality; the cultural bases to economic processes; and the hybrid assemblages of nature and the environment. What unites these new approaches is less the content of their interest than an embrace of the necessity to challenge dominant narratives with a heightened reflexivity to relationships of representation and power.

Yet however much the cultural turn has reshaped the methodological and theoretical possibilities for how research is done in human geography, many worry that the turn to culture elides or dodges key social and political concerns regarding injustice, inequality, growth, exclusion, and development. It is argued that the cultural turn relies on an excessively theoretical and not substantive—a soft—approach, which marginalizes the efficacy of investigative methods to construct a solid empirical corpus from which to make secure descriptive claims and normative or programmatic interventions. The turn has thus resulted in an overly decorative attention whose radical promises ring hollow in the face of exacerbating inequalities and injustices. Culture, it is argued, has become—ironically—simplified and is wielded as a disciplinary device, rather than being seen as the complex dynamic that matters most in our explanations and explorations.

Debates on the progress or fancy of the cultural turn expose differing visions of what humanities and social science research is meant to do. Does it change the world? Does it describe the world? Or does it do both and thereby reinvent how the world is made meaningful and pertinent to human and spatial diversity? Such debates are good for geography, which sits at the foot of unparalleled opportunity due to its inherent interdisciplinarity, unique diversities of insight and methodological approaches, and special focus on space in human and environmental relationships. Care, however,
must be exercised not to depend too heavily on cultural turn as a descriptive, disciplinary term. Overreliance risks reinforcing a narrow, historicist picture of geography. Geographical thought has always been multifaceted, contestatory, and plural. Suggesting that, as a discipline, geography made a turn to culture—or, conversely, should now limit its culturalist tenor—risks imposing, if implicitly, a common trajectory by which a less than reflexive disciplinarity might solidify its position and ignore its diverse pasts. Those on both sides of the debate would agree, however, that it is in how we explain culture, and what we do with those explanations, that the power and resonance to engage space, people, environments, and their dynamic, constituting relationships lie.

Mark Jackson

See also Critical Human Geography; Cultural Geography; Poststructuralism; Spaces of Representation/Representational Space; Spatial Turn; Symbolism and Place

Further Readings


History and Dimensions

Cyberspace should preferably not be confused with the Internet, although often the terms are used interchangeably. Electronic networks predate the Internet. Large firms, such as airline companies, started to establish electronic data interchange systems as early as the 1960s. In 1982, France pioneered the Minitel system. Although rudimentary by today standards, it was the first electronic network open to a large public.

Research on the Internet (for “inter-networks”) started in the Advanced Research Projects Agency (ARPA) of the U.S. Department of Defense. ARPANET, created in 1969, was the first packet-switching network. In the 1970s, Vinton Cerf and Robert Kahn created a new protocol for computer interconnection, TCP-IP (transfer control protocol, Internet protocol), officially adopted on January 1, 1983. Another important step, in 1991, was the creation of the World Wide Web by Tim Berners-Lee, a researcher in the European Organization for Nuclear Research in Geneva. The Web, which makes it possible to navigate between documents present on computers through hyperlinks, is the fundamental framework of cyberspace.

The growth and pervasiveness of cyberspace since the late 1990s was fuelled by the digital convergence of different media. Electronic devices, such as computers, television sets, cameras, and
mobile telephones, can share data through portable supports or networks. Moreover, a single device, such as a telephone handset, has now “multimedia” capabilities: The user may take photos, watch TV, navigate the Web, and connect to sites and applications where he or she can read newspapers, send e-mail, do shopping, play online games, and so on.

Cyberspace features a hybrid nature. On the one hand, it is dependent on material infrastructures (servers, fiber-optic cables) rooted in real geographic space. Location has therefore an actual effect on cyberspace accessibility. For example, the distance between a given house and the nearest telephone exchange has a full impact on the available DSL bandwidth. A country such as India is subject to a permanent telecommunications bottleneck due to its poor infrastructure. Cyberspace is therefore characterized by a digital divide. Urbanized regions in developed countries have high densities of networks, servers, computers, and cyberspace navigators. Peripheral regions, especially in the poorest countries, are less visible in cyberspace.

On the other hand, cyberspace can be conceptualized as a “virtual world,” where time and space seem to have collapsed. All other things being equal, given that Internet connections last milliseconds, navigation through cyberspace is almost insensitive to geography.

Since the invention of the Internet, the size of cyberspace in terms of complexity and amount of data has risen tremendously. In 2009, the Internet had more than 1.7 billion users. In 2007, the number of domain names crossed 500 million (Figure 1). Public institutions, firms, and individuals are frantically creating and posting digital content on the Web (e.g., libraries are digitizing their catalogs). Therefore, the exact size of the Web, measured in terabytes of data, cannot be measured.
with a reasonable margin of error. Pundits often make the distinction between the “surface Web”—the part of cyberspace open to the public—and the “invisible” or “deep Web,” which is not indexed and stays out of the reach of search engines. It includes, notably, companies’ intranets and the enormous amount of digital material they warehouse in data centers. In May 2008, the Web contained more than 47 billion Web pages. The “deep Web” is several times larger.

A Mirror of Society

The spectrum of human activities that use cyberspace navigation has become as large as life itself. For many applications, cyberspace is merely an electronic marketplace, a “one-stop shop” where people have access to real items and services (e.g., visiting tourist sites, booking hotels). Business-to-business electronic commerce has become a normal standard in interfirm relations Business-to-customer e-commerce lags behind: In 2007, e-commerce accounted for 3.3% of total U.S. retail sales. Individuals still prefer shopping experiences in real space rather than in cyberspace.

In many other cases, cyberspace may be conceptualized as a world in itself, although virtual, where people may spend a notable part of their life chatting, downloading movies, playing games such as “multiuser dungeons,” gambling, looking for pornography, and so on. Some social interfaces have reached a worldwide dimension. MySpace (www.myspace.com) and Facebook (www.facebook.com) are the best known. Second Life, a 3D (three-dimensional) virtual world created by Linden Research, boasts 13.6 million residents, who may participate in social events, play games, buy land and build structures, create businesses, and pay with money (the Linden dollar) withdrawn from virtual ATMs (http://secondlife.com). Addictive habits in cyberspace are common, especially among youth, sometimes ending in a social withdrawal syndrome. As a mirror of society, cyberspace has become a field for a large spectrum of illegal activities, or cybercrime: hacking, phishing, child pornography, identity theft, spying, securities fraud, and money laundering. In 2000, a study by Price-waterhouseCoopers estimated the cost to the world’s economy of computer hacking and virus writing and dissemination at $1.6 trillion.

Finally, the pervasiveness of Internet applications must not divert us from the fact that the key issues of our time (e.g., environment, energy, food, health, sustainable development) are located in real geographic space, not in cyberspace. Up to now, navigating cyberspace remains a poor experience compared with navigating the physical world.

Bruno Moriset

See also Communications Geography; Digital Divide; Information Society; Spaces of Representation/Representational Spaces; Telecommunications and Geography; Video Games, Geography and; Virtual and Immersive Environments; Virtual Geographies

Further Readings

The term *cyborg ecologies*, along with the allied notions of hybrids and nature-cultures, describes the view that the world is made up of both human and nonhuman networks. More specifically, it suggests that we must pay both empirical and theoretical attention to the connections among the biophysical, social, discursive, and technical elements of any given event, object, subject, idea, or thing.

This idea emerged in response to two different theoretical stances. First, cyborg ecologies sought to challenge the long-standing separation of nature from culture, or the notion that nature is a biophysical reality, while culture is a human construction. The concept is intended to disassemble the presumed unities of nature, culture, and technology, asserting that the world has always been made up of assemblages of animals, humans, and machines. Through this assertion, the boundaries between nature and culture, animal and human, and object and subject are unsettled. Moreover, the notion of cyborg ecologies is also on the forefront of a return to the material in the face of the so-called cultural turn in geography and its cognate disciplines, which emphasized the role of discourse in constituting the world. Thus, discourse became a primary analytical tool to explain “Nature” as a cultural container for ideas about wilderness, race, class, gender, sexuality, empire, and the like rather than a biophysical reality. While the “cultural turn” provided important insights into the way power works by making some things appear natural, it also evacuated the nonhuman from a discussion of nature, as animals and plants become yet another power effect produced through discourse. This perspective left out a whole range of nondiscursive practices, such as those of animals, plants, viruses, or technologies, for example, which, although they do not speak, can exercise their own particular forms of agency and are able to attach themselves to networks in unpredictable ways.

Using examples such as the ozone hole, nanotechnology, and mad cow disease as heuristics, ideas about cyborg ecologies insist on the recognition that the boundaries between nature and culture have always been porous and the effort to draw distinctions between these two realms is suspect. The concept of cyborg ecologies emphasizes the endless but always contingent networks among various agents such as animals, plants, discourses, technologies, and humans.

The scholarship that uses the idea of cyborg ecologies draws on the work of science studies scholars such as Donna Haraway and Bruno Latour. Although the cyborg, or cybernetic organism, has long been a feature of both science fiction and social theory, it is with Haraway’s seminal 1985 text “A Cyborg Manifesto” (expanded and reprinted in 1991 in *Simians, Cyborgs, and Women: The Reinvention of Nature*) that the term gained substantial currency in geography. In response to both the binaries found in much scholarship of the time and their manifestations in feminist thinking, Haraway borrowed the notion of the cyborg as a means to think through the interconnections between the synthetic and the organic—a playful and ironic figure that disrupts neat categories. She uses the cyborg as a political metaphor that insists on partiality, contingency and also the relationality between animals and machines. This also finds common cause with the actor network scholarship of Latour. In his book *We Have Never Been Modern* (1993), Latour asserts that modernity is made up of two interrelated processes: purification and mediation. He contends that purification is the practice of bracketing the human (or culture) from the nonhuman (or nature), but this practice can only be supported by the work of mediation, that is, the proliferation of hybrids—in other words, Haraway’s cyborgs.

In geography, one scholar closely associated with the notion of cyborg ecologies (although she never uses the term) is Sarah Whatmore. Drawing on the works of Haraway and Latour as well as Gilles Deleuze and Felix Guattari, Whatmore contends that the stale debates between nature as either pure construction or absolute reality need to be left aside in favor of an understanding of the relational character of the world, involving all manner of human and nonhuman combinations, where agency can be found in more than human intention. An example might help flesh this out. In her book *Hybrid Geographies*, she draws on the case of the leopard in Roman games (among many others) to
make the notion of cyborg ecologies more concrete. She suggests that we must view the leopard in this context as not simply a biological organism but an assemblage of the technologies, devices, instruments, programs of training, colonial discourses, practices of commerce, performances, bodies, and places that brought it from Africa to the arena, making it a recognizable life-form infused with significance and ready for its eventual slaughter. By tracing the network of the Roman leopard, from its hunting to its death, Whatmore demonstrates that while we cannot imagine the leopard as a preexisting entity, we cannot suggest that it comes into being only as a product of discourse. It is both subject and object in this cyborg ecology.

Thus, the notion of cyborg ecologies explores how we might conceive the world as it is rather than as what the boundaries between inside and outside, human and nonhuman, nature and culture, or the various other dualisms that have structured humanist thought have tried to make it. It suggests, perhaps, a way out of dualistic thinking to a more nuanced understanding of the complex human and nonhuman entanglements that have always been with us.

Stephanie Rutherford

See also Actor-Network Theory; Cultural Turn; Hybrid Geographies; Nature-Society Theory; Poststructuralism

Further Readings


Cyclones: Extratropical

Extratropical cyclones (ECs) are organized synoptic-scale low-pressure (cyclonic) weather systems that govern a considerable proportion of midlatitude (~30°–60°) weather patterns (also commonly referred to as midlatitude cyclones). ECs ultimately form in response to the global circulation and energy balance. Given that the equator-to-pole temperature gradient is greatest during the transition and cool seasons, the magnitude and frequency of ECs is also greatest during these times. Part of the Earth’s attempt to balance the surplus and deficit of net energy at the equator and the poles, respectively, involves large-scale equatorward and poleward air mass advections of relatively homogeneous thermal and moisture characteristics. The interaction along the fronts of these air masses is a key trait in the development and maintenance of ECs.

Satellites have enabled us to observe the typical comma-cloud shape of ECs, which typically have horizontal extents >1,000 kilometers, with a life cycle of several days (see photo).

The resulting mesoscale weather from a passing EC comprises a vast spectrum of events, including light rain or snow, torrential downpours, blizzards, damaging winds, destructive hail, tornadoes, as well as just overcast or clear skies. Given the areal extent, duration, frequency, and resultant sensible weather from ECs, these midlatitude weather systems often have a profound effect on many lives. Thus, our understanding of the characteristics and behavior of ECs and the role they play in weather and climate is of critical importance.

Polar Front Theory

The polar front theory (PFT) serves as the modern conceptual framework for the characteristic life cycle of ECs. The PFT, otherwise known as the Norwegian Cyclone Model, was established during World War I by the Norwegian physicists and meteorologists Jacob “Jack” and Vilhelm Bjerknes (son and father, respectively) and Halvor Solberg and the Swedish meteorologist Tor Bergeron. Together, these scientists pioneered the meteorological school of thought known as the Bergen school of meteorology, of which the PFT was the
foundation. In short, the PFT states that wave cyclones develop along frontal boundaries of contrasting air masses (e.g., continental polar and maritime tropical) and are driven by thermal advections (baroclinic instabilities) along the polar front. However, it is important to note that tropical cyclones (TCs), that is, hurricanes, may evolve into ECs. This transition commonly occurs when a TC makes landfall in the middle latitudes, losing the forcing from the latent heat source of the warm tropical ocean waters but using the baroclinic energy to help sustain the cyclonic weather system.
Extratropical Cyclogenesis

Extratropical cyclogenesis refers to the development or strengthening of a cyclone in the middle latitudes. Cyclogenesis can commence from relatively strong disturbances located in the upper troposphere or in orographic regions (i.e., leeside lows) or from amplifying frontal waves (i.e., frontogenesis). With sufficient atmospheric lift along horizontal thermal and moisture gradients, an EC commonly follows predictable stages throughout its life cycle, according to the Norwegian Cyclone Model.

Stationary Front

The presence of a stationary front sets the initial stage by providing the potential energy required for a developing EC. The stationary front separates cool, dry air to the north and warm, moist air to the south (Figure 1).

However, a middle- to upper-level disturbance (e.g., short-wave trough) superimposed along the stationary boundary is necessary for the onset of interaction between the two contrasting air masses. The key here is that while frontal instability provides the foundation for cyclogenesis, the development and maintenance of an EC depend on supportive upper-tropospheric dynamical features (e.g., the meridional extent and orientation of the upper-level long-wave trough and ridge pattern of the polar jet stream—Rossby-wave configuration, named after Carl-Gustaf Rossby).

Incipient Frontal Wave

The first indication of cyclogenesis is signified by cyclonic shear along the stationary front, which promotes a developing wave form or kink within the stationary boundary (Figure 2).

This incipient frontal wave signifies the initial call and response between the lower and upper tropospheres (i.e., a system of continuity). That is, the development of a surface low-pressure center and its resultant convergence is positioned beneath an area of upper-level divergence and just downstream of the upper-level low (i.e., the low-pressure profile is tilted upstream). If the divergent airflow in the upper troposphere exceeds the low-level convergence, the low-pressure center will likely strengthen and the EC will mature into an open wave (Figure 3).

Mature Open Wave

During this stage, the EC is marked by distinct warm- and cold-front boundaries that are anchored by a broad counterclockwise wind circulation around a low-pressure center (Figure 4).

Typically, easterly winds, amorphous cloud cover, and relatively lighter precipitation and cooler temperatures are located ahead of the warm
The advancing cold front often provides sufficient lifting of relatively warmer, moister air ahead of the cold front, often resulting in heavier precipitation events and other hazardous weather (e.g., tornadoes, hail, damaging winds). As the cold front sweeps through an area, abrupt decreases in temperature, moisture, and cloud cover are common as the troposphere stabilizes.

**Mature Occlusion**

The mature occlusion often marks the initial decay of the EC. The occluded stage is often explained with respect to a faster-moving cold front that overtakes the slower-moving warm front, forming the occluded front (Figure 5).

A more accurate description of this stage involves the orientation of the low-pressure profile. The occlusion takes place when the surface low becomes positioned beneath the upper-level low and cooler air. This rearrangement of the low-pressure profile is what is referred to as a *stacked low*. When a low becomes stacked, the upper-level support becomes less conducive for maintaining the EC life cycle (i.e., a considerable decrease in baroclinic instability). However, with a supportive upstream Rossby-wave configuration, the cyclogenetic cycle can remain active and spawn what is called a “family of cyclones.”

**Conclusion**

The PFT is a simplistic conceptual model for explaining the development and evolution of an extratropical cyclone and remains widely accepted in our current understanding of ECs.
Occluded cyclones represent the late-mature stage of short-lived, rapidly moving extratropical cyclones that, with much slower-moving anticyclones, are responsible for the day-to-day weather changes in the middle latitudes. In redistributing energy and moisture, occluded cyclones are an integral component of the general circulation of the atmosphere. Juxtaposition of contrasting air masses with different temperatures and densities that resist mixing promotes the establishment of a front. A wave development on this front initiates cyclonic circulation and the genesis of a low-pressure region. The front develops discrete warm and cold portions separated by a warm sector. Through time, the cold front moves more rapidly than the warm front and eventually overtakes the warm one. Then, occlusion takes place. Precipitation associated with occluded cyclones may be of moderate to heavy intensity, resulting in significant amounts of rainfall, especially during the partially occluded phase.

Jet Streams and Surface Weather

In the midlatitudes, the general airflow pattern is from west to east, termed the Westerlies, and this dominates throughout the depth of the troposphere. At upper-troposphere levels, the airflow commonly resembles a series of horizontal wave-like oscillations, referred to as Rossby waves, that encompass most of the midlatitude region. The nature of airflow through an upper-level Rossby wave has important ramifications for the development of weather systems at the surface.
The fastest-flowing airflow usually occurs in the trough of a Rossby wave, but as the air exits the trough, changes in the flow along the upper wave lead to divergence and convergence. When the air in the upper troposphere diverges (spreads out), this draws air upward from below and encourages the surface genesis of a low-pressure center accompanied by cloud development and precipitation. A zone of divergence exists just downwind of an upper-level trough (Figure 1). In contrast, immediately downwind of an upper-level ridge, convergence promotes subsidence, which inhibits cloud development and induces pressure to rise at the surface and anticyclone formation to occur.

Low-pressure development at the surface originates with wave development on the front that separates two air masses of markedly different temperature and humidity. Cyclogenesis continues through stages to produce a mature extratropical cyclone—a synoptic-scale feature some 1,500 to 3,000 kilometers in diameter.

**Frontal Development From Maturity to Occlusion**

The mature extratropical cyclone comprises a distinct low-pressure region at the crest of the wave. A warm front occupies the leading edge of the wave that normally progresses east or northeastward. The warm front is followed by a warm sector comprising the warm air enclosed by the wave on its southern side. Finally, the cold front leads the cold air mass at the rear of the wave.

Figure 2 illustrates how, as the extratropical cyclone continues to deepen, the cold front moves faster relative to the warm front and the area of...
Figure 2  The process of occlusion
Source: Author.
the warm sector contracts. At the warm front, the warm lighter air is moving forward and upward away from the surface over the heavier colder air with which it comes into contact. It can also be seen that the cold air at the rear of the cyclone is undercutting the warm air and pushing it forward and upward at the cold front. At both the warm and the cold fronts, cooling and condensation occur within the rising air associated with the warm conveyor belt, inducing cloud development and zones of precipitation in advance of the fronts. Eventually, an occluded front forms where the cold front overtakes the warm front, with the warm air being completely lifted off the ground. This ascent and cooling of air in the vicinity of the occluded front generates precipitation.

Occluded fronts, unlike warm and cold fronts, do not separate tropical from polar air masses. Instead, at the surface they appear as the boundary between two polar air masses. A warm air mass exists aloft but is detached from the surface. Occlusions are classified as either warm or cold, depending on the relative states of the air masses lying in front and to the rear of the warm sector. A cold occlusion occurs when the air behind the original cold front is colder than that ahead of the warm front. In the reverse situation, it is termed a warm occlusion, and indeed, most occlusions are warm. Air in advance of the cyclone is likely to be coldest when cyclones occlude over Europe in winter and very cold continental polar air is affecting the continent. Similar circumstances are observed in warm occlusions over British Columbia and the Pacific Northwest of the United States when the relatively mild maritime polar air behind a Pacific cold front moves onshore and encounters the colder polar continental air.

The extent of an occluded front progressively increases away from the center of the cyclone and, in so doing, eliminates the waveform at the surface. Eventually, the cyclone dissipates, but often with the polar front becoming reestablished further south than previously. Extratropical cyclones usually occur sequentially in families, with each new system forming progressively south on the trailing cold front of its predecessor. Eventually, the front trails far to the south, and frontolysis (frontal decay) occurs when there cease to be differences in the temperature and humidity properties of adjacent air masses.

By no means do all frontal cyclones follow the sequential development outlined above, and not all extratropical low-pressure systems originate as frontal waves. This is more generally characteristic of oceanic cyclogenesis, and over Central North America, cyclones forming in winter and spring depart considerably from the standard pattern.

**Case Study of an Occluded Cyclone**

While the process of occlusion is a later stage of the life cycle of an extratropical cyclone, significant and intense precipitation may be associated with the occluded front. This is exemplified by the extratropical cyclone of October 21 to 22, 1987, which produced significant rainfall totals categorized as “remarkable” and associated record floods in Northern Ireland.

**Synoptic Situation, October 20 to 22, 1987**

At the 500-hPa (hectopascal) level, a marked cold trough located west of Ireland at 12:00 UTC (Universal Coordinated Time) on October 20 moved east over the next 24 hrs. (hours), with a closed low developing over Ireland, apparent on the 500-hPa chart for midday on October 21 (Figure 3a). At 00:00 UTC on October 21, a shallow cyclone (LD) with surface pressure of 1,001 hPa was moving north-northwest. An associated warm front, aligned north-south along the west coast of Scotland, and through the Welsh Borderland, was moving slowly westward (Figure 3b). By 06:00 UTC on October 21, the deepening cyclone had moved up the Irish Sea; a distinct westward moving warm front and a northward progressing cold front then affected Northern Britain (Figure 3c). Further deepening occurred with continued northward movement, and at noon, the cyclone was centered over the North Channel (Figure 3d). The cold front moved rapidly, turned cyclonically, and by early afternoon of October 21, an occlusion had formed over Northern Ireland, which affected the province during the remainder of the day. By 00:00 UTC on October 22, the cyclone lay off Western Scotland, and the associated chart (Figure 3e) indicated the rotational aspect of the system, with the occluded front in the southern sector of the cyclone returning eastward and skirting the north coast of Ireland.
Figure 3  Synoptic situation, October 20 to 22, 1987: (a) 500-hPa height (dam, solid lines) and 1,000- or 500-hPa thickness (dam, dashed lines) for 12:00 UTC on October 21; (b) to (f) surface charts for 00:00 UTC, October 21, and 00:00 UTC, October 22, with isobars at 4-hPa intervals

Source: Adapted from Daily Weather Summary, London Weather Centre, Meteorological Office.
Precipitation began in the south of the province before 01:50 UTC on October 21 and by 03:00 UTC in the west; and by 04:45 UTC, precipitation was experienced throughout Northern Ireland. This lasted 20 to 25 hrs.

Upland sites generally recorded twice the precipitation amount as adjacent lowlands. This orographic enhancement resulted from differential intensities rather than durations. Most upland sites recorded over 75 mm (millimeters), the highest areas more than 100 mm (Figure 4). The greatest total (137 mm) occurred at Glenanne, County Armagh, while at many sites, amounts were categorized as “very rare,” with return periods of 160 years or more. Precipitation totals of less than 50 mm were confined to the east coast, the north coast fringe, and an area extending across part of the mid Antrim Plateau.

**Figure 4** Total precipitation (in millimeters) between 00:00 UTC, October 21, and 09:00 UTC, October 22, 1987

*Source:* Adapted from data provided by the Meteorological Office, Belfast, Northern Ireland.

**Total Precipitation Distribution**

Precipitation began in the south of the province before 01:50 UTC on October 21 and by 03:00 UTC in the west; and by 04:45 UTC, precipitation was experienced throughout Northern Ireland. This lasted 20 to 25 hrs.

Upland sites generally recorded twice the precipitation amount as adjacent lowlands. This orographic enhancement resulted from differential intensities rather than durations. Most upland sites recorded over 75 mm (millimeters), the highest areas more than 100 mm (Figure 4). The greatest total (137 mm) occurred at Glenanne, County Armagh, while at many sites, amounts were categorized as “very rare,” with return periods of 160 years or more. Precipitation totals of less than 50 mm were confined to the east coast, the north coast fringe, and an area extending across part of the mid Antrim Plateau.

**Distribution of Shorter-Period Precipitation Amounts**

Prior to occlusion, as the synoptic system encroached northward between 00:00 and 06:00 UTC on October 21, a major precipitation area developed, producing more than 15 mm over the
Between 06:00 and 12:00 UTC, as the cyclone and its associated fronts crossed eastern areas of the province, 6-hr. totals were recorded of 40 mm in the Mournes, 35 mm over upland South Armagh, 25 mm over areas of the Antrim Plateau, and 20 mm throughout the Sperrins.

As the cyclone began to occlude after midday, between 12:00 and 15:00 UTC, a distinctive feature of precipitation receipt was a relatively dry zone in the east of the province (Figure 5a). The west and south remained wet, but the greatest totals now occurred along the north coast, and Coleraine received 13 mm between 12:00 and 13:00 UTC. This marked contrast of wetness across the region coincided with the occluded front having crossed eastern areas and having become aligned from Londonderry to the Mournes.

A return to precipitation of moderate to heavy intensity in eastern areas between 15:00 and 18:00 UTC occurred, with cyclonic rotation of

---

**Figure 5**  Surface precipitation (in millimeters) that fell in the period (a) October 21, 12:00 to 15:00 UTC; (b) October 21, 15:00 to 18:00 UTC; (c) October 21, 18:00 to 21:00 UTC; and (d) October 21, 21:00 to 00:00 UTC, in 1987

the occluded front and the edging back eastward of associated rain bands (Figure 5b). The Belfast area consequently experienced intensities of up to 8 mm/hr., and a drier zone replaced the rain band that had affected the northernmost areas in the previous 3 hrs.

The precipitation patterns for the period October 21 18:00 to 21:00 UTC (Figure 5c), and the subsequent period of 3 hrs. to 00:00 UTC on October 22 (Figure 5d), indicate that the linear rain band extending over the eastern part of the province was aligned parallel to the occluded front. Furthermore, as the depression and the associated rotating occlusion progressed north-northwest, the precipitation signatures associated with this precipitation area depicted the position of the front.

After 00:00 UTC on October 22, the southern areas became virtually dry. The only significant precipitation occurred around Londonderry, along the north coast and over the northernmost areas of the Antrim Plateau as the occluded front skirted the north coast.

The heavy precipitation of October 21, 1987, was a significant event in the climatological history of the province. Widespread—and in places, severe—flooding occurred throughout Northern Ireland. The physical, social, and economic repercussions had impacts on both urban and rural communities, with Strabane, Omagh, Lisnaskea, and the Coleraine area among the most affected localities. Furthermore, the event demonstrates how occluded cyclones can impart significant weather before progressing to the dissipating stage of the life cycle of extratropical cyclones.

Nicholas L. Betts

See also Air Masses; Atmospheric Circulation; Atmospheric Moisture; Climate: Midlatitude, Mild; Climate: Midlatitude, Severe; Cyclones: Extratropical; Floods; Hurricanes, Physical Geography of; Precipitation Formation; Weather and Climate Controls

Further Readings

DANGERMOND, JACK

Jack Dangermond and his wife Laura are the founders of the Environmental Systems Research Institute (ESRI), the world’s fourth largest privately held software company. Founded in 1969 and headquartered in Redlands, California, ESRI is arguably the technical and market leader in geographic information system (GIS) software. ESRI is the largest commercial GIS software provider in the world, with more than 1 million users in more than 100,000 organizations representing government and nongovernmental organizations, academia, and industries such as utilities, health care, transportation, telecommunications, homeland security, military, retail, and agriculture industries.

Dangermond initially founded ESRI as a small consulting firm/research group to perform land use analysis. However, its focus evolved into GIS software development, highlighted by the release of ARC/INFO in the early 1980s. The development and marketing of ARC/INFO positioned ESRI with the dominant market share among GIS software developers. Today, ESRI is the largest GIS software developer in the world, and its core product, ArcGIS, has direct linkages to Dangermond’s initial efforts in developing ARC/INFO during the early years of the company’s inception. Workstation ARC/INFO evolved into ESRI’s flagship product, developed for spatial analysis, now aptly called ArcGIS, which is a desktop graphical user interface (GUI)-driven product that possesses myriad spatial analysis functionalities.

Dangermond graduated with a bachelor of science degree in environmental science from California State Polytechnic University in Pomona, California. He holds a master of science degree in urban planning from the Institute of Technology at the University of Minnesota and a master of science degree in landscape architecture from the Graduate School of Design, Harvard University, where he worked in the Laboratory for Computer Graphics and Spatial Design. The experience there led to the eventual creation of ESRI and its subsequent software products. Dangermond also holds several honorary doctorates from various higher-education institutions around the globe.

Dangermond fostered the growth of ESRI from a small research group to an organization of more than 2,700 employees, known internationally for GIS software development, training, and services. In 2007, ESRI had 16 subsidiaries and 72 distributors worldwide. ESRI also has 11 regional offices throughout the United States. Dangermond is recognized not only as a pioneer in spatial analysis methods but also as one of the most influential people in the realm of GIS. Dangermond explains why the development of GIS techniques and technologies has become such an integral part of his vision pertaining to geographic research and its role in society: “We are visual creatures. Seeing promotes understanding. If GIS
helps us see what is happening in a larger, graphical context; and helps us to model and estimate change; perhaps more collaboration and compromise will result.”

If the discipline of geography is the framework for understanding, conceptualizing, modeling, and visualizing the world, then GIS is a tool that facilitates and enhances the process of integrating what we know. Dangermond has arguably built a career and a company that endeavor to accomplish this task.

Shawn Lewers

See also Database Management Systems; Geospatial Industry; GIScience; GIS Software

Further Readings


DARBY, HENRY CLIFFORD

(1909–1992)

Henry Clifford Darby, arguably Britain’s best-known historical geographer, sought to break down the barriers between history and geography as separate disciplines. He produced a prodigious body of scholarship, especially concerning medieval England.

Born in Wales, he began attending Cambridge University at age 16 and then received a teaching post there on graduation with the PhD, soon becoming a college Fellow. During World War II, he worked for the Naval Intelligence Division of the British Admiralty and then left Cambridge to assume the John Ranking Chair at the University of Liverpool from 1945 to 1949. He then relocated to University College London before returning to Cambridge to hold a chair of the geography department in 1966. He retired in 1976 and was knighted in 1988. Intellectually, he was greatly influenced by the works of Paul Vidal de la Blache and Lucien Febvre.

Darby wrote extensively on the medieval history and geography of England, including topics such as the draining of the English fens, about which he published two books in 1940. (While most of his work focused on England, he also coauthored books on the history of France and Greece—in 1942, he produced a 1,630-page handbook on France.) He emphasized the cross-sectional or horizontal approach to historical geography, that is, the reconstruction of past geographies at a given moment in time, as opposed to the vertical approach, which centers on historical change. His focus remained relentlessly on landscapes: In his words, he was not much interested in shifting from “places changed by men to men as changers of places.” In this approach, Darby stood in contrast to contemporaries such as Carl Sauer (whose work he knew), who advocated a more temporally dynamic perspective of landscape change.

By far his most well-known work, one that consumed much of the latter part of his life, consisted of a gargantuan seven-volume tome on the Domesday geography of England written between 1952 and 1977, based on the famous Domesday Book of 1086 (also known as the Book of Winchester) following the Norman conquest of 1066. (The title Domesday comes from the Old English dom, an account or reckoning, which is also the source of the word doom.) William the Conquerer, eager to understand the new lands he had acquired, commissioned something like a medieval census, largely for purposes of taxation. The Domesday Book consists of detail inventories of people, farmland, crops, animals, forest cover, and other phenomena, providing an enormous trove of information about medieval geography. The entire project appears to have been carried out by one person. Darby’s work carefully translated, summarized, and mapped this information and noted its historical context. Subsequently, much of this information has been digitized.

Darby’s work and scholarly views were largely traditional and empirical, as was common with the subdiscipline of historical geography before the rise of social theory in the late 20th century. Darby himself recognized that the discipline had changed irrevocably in his time. Nonetheless, he had considerable influence on subsequent historical geographers, and his work continues to draw
admiration for its careful attention to detail and painstaking reconstruction of past landscapes.

Barney Warf

See also Historical Geography

Further Readings


Although Charles Darwin was actively engaged in geographical exploration and his work would have been hardly possible if not for geographical differences in the distributions of species and characteristics, Darwin has received little direct attention in geography. Despite Darwin’s lack of recognition as an influential figure for geography, much geographical research since the publication in 1859 of On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life has been either explicitly or implicitly inspired by Darwin. David Stoddard suggested that the Darwinian theme of evolution as change through time, with emphasis on the relationship between organic life and the environment, and his analysis of selection and struggle, exerted a strong influence on research in geomorphology, pedology, and ecology; led to the organicist interpretation of regions and states; and resulted in a deterministic application of the concepts of selection and struggle in human and political geography. Another major theme of Darwin’s work, the random nature of original variations, was ignored by geographers until the late 20th century, which was largely the result of Darwin’s inability to provide a good explanation for it.

The impact of Darwin’s work constituted nothing short of a revolution in science, confronting the old mode of doing natural history head-on. Rather than attributing the diversity of natural forms to the will of a divine creator, Darwin’s book provided a theoretical framework that explained the diversity of species in a purely naturalistic, causal fashion. Darwinism itself is best understood as an evolving entity as opposed to a homogeneous body of knowledge. If the interpretations and meanings of Darwinism change and evolve, can we still identify a set of conceptual pillars that make up Darwinian evolutionary theory? At a general level, we can identify a set of concepts and principles that characterize a Darwinian view of evolution:

- At any moment in time there is sustained variation of species produced through random mutation.
- The environment winnows on this variation through natural selection. Individuals and species that are relatively better adapted to their environment are more likely to leave more offspring and so expand at the expense of relatively poorly adapted individuals and species. The competition for scarce resources does not require “survival of the fittest.”
- There must be a mechanism that keeps variation stable long enough for selection to operate. If species were to adapt instantaneously to demands of their environments there would be no variety left to select from. Although Darwin realized that there must be a mechanism to pass on information between generations, he was never able to prove that, because genetics was developed only after the publication of The Origin of Species.
- There is no progression or development toward some ideal, perfect, or higher stage of development. Although humans are more complex than single-cell organisms, they are not necessarily better in terms of some preconceived ideal and optimal state of adaptation. They simply fill different niches in an environment.

Some writers have argued that Darwin’s evolutionary principles can be generalized, although the mechanisms by which variation is produced, maintained, and destroyed are domain specific. While
random mutation is the only source of variation in biological systems, a deliberate search for improvement and problem solving generates variety in social systems. While genes pass on information from one generation to the next in the biological domain, learning and knowledge acquisition are the primary means of passing on information in social systems. Collective knowledge is often embedded in institutions and organizations, constraining and enabling the creation, transmission, and destruction of new knowledge. Although natural selection produces gradual evolution through the elimination of relatively ill-adapted forms and differential growth of relatively well-adapted forms, there is no reason to assume that competition is always beneficial for the survival of species. In various biological and social contexts, altruistic and cooperative behavior may benefit some groups over others in the evolutionary process.

**Evolution and Time**

In his work on the development of landforms, William Morris Davis explicitly referred to a cycle of life and used terms such as *birth, youth, adolescence, maturity, old age, second childhood, infantile features,* and *struggle* to emphasize the analogy of an organism undergoing a sequence of changes in form through time. Landscapes rose through tectonic uplift and were then shaped by erosive agents until they reached the end of their life cycle, a low plain. The idea of evolution as change through time also found its way into plant geography, social anthropology, the historical interpretation of technological development, or the understanding of cities as organisms changing through time. In geography, evolution soon implied little more than the idea of change, development, and progress, the simple linear and irreversible development of something over time, rather than the outcome of a process characterized by the interplay of variety, selection, and inheritance.

The geological metaphor of layering and evolution as change was picked up late in the 20th century again by economic geographers to conceptualize economic landscapes as layers of accumulation where value sunk in fixed capital gradually builds up over time. The basic idea of value accumulation driven by the logic of capital and class struggle comes from Marxist economic geography. However, the fact that economic landscapes change slowly and that sunk capital, industrial structure, capital-labor relations, and institutional environments constitute relatively stable entities that transfer information over time and form barriers to instantaneous adaptation to new economic requirements fits squarely with gradual evolutionary change. While existing regional value configurations represent adaptations to the past or current requirements of the economic environment, they are often difficult to adapt to future requirements. Therefore, further change often requires the violent destruction of previously well-adapted regional value configurations through plant closures and relocations, job loss, and decay of economic and social infrastructure.

**Organization and Ecology**

Darwin was impressed by the complex interrelationships and connections between living things and their environment, giving rise to organicist interpretations of entities above the level of the individual. In geography the organism analogy operated at the level of the Earth, its regions, and the state, and predates Darwinian evolutionary theory, although much of this work was teleological. For Alexander Von Humboldt and Carl Ritter, unity, harmony, and independence of the parts in nature were the organic analogy that they interpreted as evidence of the work of a divine creator. In regarding regions as organic wholes, French and German regional geographers of the 1920 and 1930s reached conclusions similar to those of Humboldt and Ritter, writing half a century earlier.

The organism analogy was picked up by Friedrich Ratzel in his *Anthropogeographie* and “The Territorial Growth of States,” where he interprets states as expanding organisms with an increasing need for resources and living space (*lebensraum*) such that geographic expansion is inevitable. Ratzel not only adopted the organism analogy but was also influenced by Ernst Haeckel’s ecology, economist Herbert Spencer’s view on survival of the fittest, and Jean-Baptiste Lamarck’s views of evolution allowing for the direct and immediate influence of the environment on human behavior. For Ratzel, there existed a
strong, indivisible, and almost mystical relationship between people and land. The organic quality of states depended on the organization and interdependence of the parts. But by adding the properties of growth and competition, Ratzel goes beyond the organic analogies of the Earth and geographical regions, as proposed by Paul Vidal de la Blache and Alfred Hettner.

In the early 1900s, science still struggled with the relationship between individuals and wholes, and the organic analogy was little more than a metaphor hinging on formal and functional comparisons between living matter and complexly interrelated facts in areas rather than causal explanations. Because the organic analogy was unable to guide empirical investigation, it was largely dismissed during geography’s quantitative revolution of the 1950s and 1960s, in which it was transformed from an idiographic to a nomothetic science promoting methodological reductionism in the form of either methodological individualism or spatial fetishism. Recent developments in complexity science may finally put the part-whole dilemma to rest and may provide new theoretical foundations for an organicist geography.

**Selection and Struggle**

A critical argument in Darwin’s theory is natural selection and the struggle for existence. Environments shape species through their influence on reproductive success. Relatively better-adapted species are more likely to survive and reproduce, such that they expand their size in the overall population relative to poorly adapted ones. Darwin’s theory explaining adaptation in nature by variation and natural selection was based on empirical evidence, but he could not, before the discovery of Gregor Mendel’s work on genetics, offer any explanation of the origin and persistence of variation. And although Darwin’s evolutionary theory is based on probabilistic principles where random variation is probabilistically translated into differential growth of different variants, Darwin himself never used the word probability, even though Pierre-Simon Laplace’s work on probability was already available at the beginning of the 19th century. This limitation of Darwin’s work made room for Lamarckian evolutionary theory, where species adapt to environmental conditions and where those changes are then passed on to the next generation. In this view, the environment is deterministically related to the behavior of individuals and so opens up the door to environmental determinism in geography.

While the organism analogy simply indicated that regions or states could be considered as wholes that are more than the sum of the individual parts, the struggle for existence was now applied to the struggle of states for living space. States, like other organisms, have a natural tendency to grow and expand. With increasing population size, the demands to feed the population increase, and so the struggle for territory increases. The size of the territory of a state is then the spatial expression of the level of civilization. Ratzel’s theory was in line with German imperial thinking at its time as it provided a “scientific” underpinning for its expansionist policy. The same lines of thinking also provided the underpinnings for British and American empire building. For Halford Mackinder, the key to British survival was control of the “Heartland,” the vast area of land in Eurasia immune to sea attacks, where civilizations could create a natural power base. The key for British policy was then to make sure that no single power would control this mass of land, from where it could develop a power base to control the world.

While Ratzel’s struggle for lebensraum is a crude interpretation of Darwin’s struggle for existence, the question of who wins the struggle is a different one. It was politically expedient to find a natural explanation and justification for European and U.S. dominance at the beginning of the 19th century, and a crude Lamarckian interpretation of the relationship between the environment and people would deliver just that. Environmental determinists argued that climate or frontier conditions were selected directly for specific human traits and influenced ideas, racial characteristics, and cultures. In the crudest interpretations, differences in climates would so account for differences in intelligence and political and economic power. The writings of Ellen Churchill Semple, Ellsworth Huntington, and Frederick Jackson Turner are examples of these lines of thinking.

Russian geographer Pyotr Kropotkin advanced the notion that the struggle for existence does not necessarily entail competition between individuals,
DARWINISM AND GEOGRAPHY

regions, and states. In his travels to Siberia, he was impressed by the cooperation among species to survive in harsh environments. Altruistic and cooperative behavior could thus favor groups of individuals better than mere competition. While the empire and frontier politics of Europe and the United States highlighted competition as the main force in the struggle for survival, socialist governments emphasized cooperation and altruism.

Recent Developments

The racial politics of the German National Socialist Party brought the application of all biological theories, concepts, analogies, and metaphors in the social sciences into disrepute. The demise of biological concepts in the social sciences was further accelerated through the propagation of methodological individualism in sociology, economics, and political science (particularly in the United Kingdom and the United States), leaving little room for organicist ontologies.

It was only in the 1970s that a few economists started to challenge the neoclassical pillars of the perfectly rational individual and competitive equilibrium. In their attempt to overhaul mainstream economics, they looked for inspiration in biology rather than mechanical physics and so reintroduced the Darwinian notions of variety, heredity, selection, disequilibrium, and human-environment interdependencies into economics. Similarly, network metaphors and notions of embeddedness started to resurface in sociology.

Evolutionary economics was picked up by economic geographers in the 1990s to account for regional differences in technology evolution and institutional environment and to offer a dynamic explanation of the uneven evolution of the space economy. Although some of this work draws its inspiration from Darwin, others prefer complexity theory or the principle of path dependence as theoretical anchors for the development of an evolutionary economic geography. The organicist ontology is increasingly invoked to challenge methodological individualism and highlight the irreducibility of wholes to parts. This opens up the possibility of emergent properties at various spatial scales, with spatially varying features such as culture or institutions exerting a downward causal pressure on individual behavior.

With Darwinism and complexity science evolving at a rapid pace, these newer applications are more careful in their transposition of causal processes and concepts from the biological to the social domain. While a general theory of evolution based on the principles of variety, inheritance, and selection may hold in different domains, domain-specific auxiliary concepts and hypotheses are required to establish how variety is produced, maintained, and destroyed differently in different domains and what varieties of characteristics matter for the evolution of environments, species, humans, institutions, or regions.

Jürgen Essletzbichler

See also Biodiversity; Ecosystems; Environmental Determinism; Geopolitics; Nature; Ratzel, Friedrich; Semple, Ellen Churchill; Social Darwinism

Further Readings

Dasymetric mapping is a cartographic technique with roots in choropleth mapping. Both techniques originated in the 1800s when cartographers and others began to design maps of population distributions. Like the choropleth map, the dasymetric map uses area symbols to show variations in topics collected by standard political units (e.g., census tracts, county boundaries). The advantage of this particular representation lies in its treatment of the boundaries of the political units. A choropleth map of county population density, for example, would suggest uniform densities within each county, with possible abrupt changes occurring between adjacent counties due to the geography’s artificial nature.

The mapping of the topic with this technique is heavily influenced, then, by the geography of the political units used; displaying population densities by census tracts may result in quite different density patterns when compared with a similar map using counties as the geographic foundation (Figure 1).

The dasymetric version of the same data uses ancillary data sets, such as land use, as the boundaries for representing changes in population density (Figure 2).

The density data from each county are reassigned to land use areas within the county using formulas to determine how many people should be assigned to each inner-county area. The result is a more accurate mapping of the topic; areas of abrupt change follow more closely the real-world distribution because they are tied not to political

**Figure 1** Population density of Guilford County, North Carolina, by census tract and block group

_Source: Map created by author based on data from ESRI Data and Maps, 2000; Census 2000 Summary File 1 (SF 1) 100-Percent Data, U.S. Bureau of the Census, 2000; and National Land Cover Database, Multi-Resolution Land Characteristics Consortium, 2001._
boundaries but to natural boundaries. The initial interest in dasymetric mapping seems to have originally peaked during the 1920s and 1930s, around the time that J. K. Wright published his seminal work on the subject. It is back at the forefront of cartographic research today, largely because of its more precise representation potential and its ease of implementation with current state-of-the-art software.

Elisabeth S. Nelson

Figure 2  Population density of Guilford County, North Carolina, using dasymetric mapping


Further Readings

DATABASE MANAGEMENT SYSTEMS

Developed in the 1960s by computer scientists who required an efficient storage and access mechanism for large data sets, database management systems (DBMS) have become a core component of present-day information systems. Database technology drives billions of transactions each day, from grocery store purchases to Web searches. The technology is also critical to research applications. DBMS facilitate efficient, scalable, secure storage of data and include query and manipulation tools that can be controlled via the standard structured query language (SQL). Properly applied database design techniques produce databases that eliminate data redundancy and provide for efficient queries across multiple tables. In geography, DBMS are used in multiple ways but are best known as a key component of geographic information systems (GIS), where they enable the storage and manipulation of attribute data associated with spatial features. Certain GIS data formats contain file-based attribute data, while others store attributes in a database table. Spatial databases are specially constructed to hold spatial features and their attributes and to provide functions for performing spatial analysis operations within a database.

Definition and History of Database Models

A database is simply a structured collection of related data. While the term database could refer to any collection of data, such as a set of alphabetically sorted recipes on index cards, it most commonly refers to a set of data stored electronically. A DBMS consists of a database and the associated software constructs that not only maintain interrelationships among the stored data but also provide methods for data insertion, manipulation, and extraction.

A database model describes the structure of a database. Several DBMS implementations emerged in the 1960s that employed either the network model or the hierarchical model, which together represented the first generation of database models. Although these early models differed in implementation, they both represented data as records with links (from one record to another) that described relationships among the records. This approach provided limited search capabilities and was inefficient when dealing with databases containing empty records. In 1970 Edgar Codd described the relational model, which represented a groundbreaking approach to database modeling that solved many limitations of the linked-list approach of the first-generation models. Codd’s work eventually guided the development of the relational DBMS, which remains the most widely implemented DBMS standard today. In later years, the object-relational model emerged, which supported the object-oriented programming concepts of objects, classes, and inheritance within the database. Many major DBMS support both relational and object-relational models.

Relational Database Management Systems

Relational databases offer many advantages over file-based information systems. Although text files are simple, are easy to read and edit, and can be processed in many ways, they do not scale well when dealing with large data sets, have limited structuring, and may contain redundant data. A relational DBMS can handle enormous data sets and maintain data integrity, eliminate data redundancy, enable user-based security, and provide structured methods for querying and manipulating the data.

Relational Database Structure and Design

A relational database contains a set of tables of related data. A table is a mathematical relation that contains a set of tuples (rows) and attributes (columns). Each table cell must contain a single value corresponding to the data type assigned to
the attribute (i.e., the value “541” would be allowed for an attribute of an integer data type, but the value “541, 503” would not). Rows can be arranged in any order within the table. Each row must be unique; data may be duplicated across rows for some attributes but not for all. To ensure uniqueness, an identifier called a primary key that holds a value unique to each row is often defined in each table. A single column (e.g., the state_id column in Figure 1) or multiple columns together can form a primary key.

A column or set of columns that match the primary key of another table may be designated as a foreign key. Tables can be linked on common attributes by defining relationships between primary and foreign keys, and the use of foreign keys enables referential integrity, an important aspect of relational databases. Referential integrity is enforced by the relational model such that any foreign key value must have a matching primary key value in the related table (as in Figure 2). This ensures that the relationships between tables remain valid by guarding against accidental deletion of important data or the creation of “orphan” records that contain old data. Primary and foreign keys are also used to establish relationships (one-to-one, one-to-many, many-to-many) between tables, which helps prevent the insertion of redundant data. A classroom of students is an example of a one-to-many relationship (one classroom has many students). Figure 2 illustrates the modeling of a one-to-many relationship.

The structure of a relational database determines its usefulness in transforming data into meaningful output. Proper database design requires the consideration of many factors, including data types, normalization, relationships, referential integrity, and common queries that will be performed. The needs of the project represented by the database must be determined prior to the development of the database, and careful planning must take place to identify how the database design will meet those needs. An important aspect of the design process is database normalization, which involves breaking down tables until they represent single entities and their columns describe only those entities. A fully normalized database has no redundant data across tables. Graphical database modeling software tools can aid in database design, including the modeling of table relationships. Many such tools can automatically generate SQL statements for creating the tables and relationships depicted in the design.

**Query Language and Methods**

Modern relational DBMS provide for query and manipulation of data via SQL. Although SQL is designated as both an ANSI and ISO standard, variations in its implementation exist across different relational DBMS, so that the syntax of a query in Oracle might vary slightly from the same query in PostgreSQL. The SELECT command forms the basis of all SQL queries and can be
modified by several clauses, including FROM, WHERE, and ORDER BY to extract the desired information from a table (or set of tables). Common functions that handle tasks such as date conversions can be placed inside the query as well. For example, in a PostgreSQL table (such as Figure 1), the following SQL command would return a list of all states established after the year 1850, in ascending order by date:

```
SELECT name, est_date FROM States
WHERE to_char(est_date, 'YYYY') > 1850
ORDER BY est_date ASC;
```

These commands enable virtually all data input and modification operations within the relational DBMS. As an example, assume that “Oklahoma” was misspelled as “Oklohoma” in the name column of the States table (Figure 1). The following SQL command would fix all instances of the error (in this case, a single instance):

```
UPDATE States SET name = 'Oklahoma' WHERE name = 'Oklohoma'
```

Relational DBMS provide a command-line interface for directly submitting SQL commands to the database. Many vendors also create graphical user interfaces for connecting to the database that can make certain query and management tasks more efficient. Programmatic access is yet another method for data access and manipulation—many programming languages such as Perl, PHP, Python, Java, and C++ provide built-in functions for connecting to a remote database and performing queries or data manipulation tasks. This powerful access method enables much of the behind-the-scenes database transactions that take place every day. A great deal of Web sites are “database driven,” with active processes performing database

Figure 2  A representation of two relational database tables from a logical data model. The tables possess a one-to-many relationship through the common state_id column, designated as a foreign key in the Cities table.

Source: Author.
operations on the back end, and Web server code pulling data from the database to the front end, dynamically populating the browser with content.

**Database Management Systems in Geography**

In geography, as in other disciplines, DBMS have been used for many years as general tools for storing and accessing diverse data sets; however, the most prominent use of DBMSs in geography today is within the context of geographic information systems (GIS). A GIS is an information and analysis system that links data attributes with spatial features by geographic coordinates. As such, a GIS contains not only tools for working with spatial features but also a DBMS for handling attribute data.

**GIS Data Formats**

Early efforts at encoding geographic data in large databases were pursued by the Canadian Geographic Information System (CAGIS) and by the U.S. Census Bureau, which developed the Dual Independent Map Encoding (DIME) scheme in the late 1960s to prepare for automated geocoding of the 1970 census. By 1990, the Bureau had integrated all their spatial data into a database called TIGER (Topologically Integrated Geographic Encoding and Referencing), an important data product that contains spatial feature and attribute data for the entire United States.

Early GIS data formats consisted of an attribute data file associated with a spatial features file by a common identifier. Some of these formats remain popular today. For example, the ESRI shapefile format uses the standard dBase .dbf file format to store attributes—a .dbf file is linked in a one-to-one relationship with a matching .shp file (which stores the spatial features) by a .shx index file. Together the three files constitute a single shapefile. Many GIS software programs can read shapefiles and have DBMS tools for manipulating attribute data stored in the .dbf format. The proprietary ESRI ARC/INFO coverage format, which predated the shapefile format, also uses a file-based approach.

More recently, ESRI introduced the geodatabase format, in which spatial features and attribute data are stored together as a hierarchical group of data objects. ESRI’s “personal geodatabase” stores data objects within a Microsoft Access database, while the “file geodatabase” stores data in a flat file structure. Enterprise versions of the ESRI geodatabase allow the use of more advanced relational DBMS technologies such as Oracle, SQL Server, or PostgreSQL with the geodatabase framework.

**Spatial Databases**

Relational DBMS products such as Oracle Spatial or PostgreSQL with the PostGIS extension add geometry data types that enable the storage of spatial data features (points, lines, and polygons) and also add spatial indexing schemes to facilitate fast searching across spatial features. These additions essentially “spatially enable” tables, allowing the database to not only store and manipulate spatial data and their attributes but also perform advanced spatial analysis operations traditionally done using a desktop-based GIS. For example, using a PostgreSQL/PostGIS table representing city limit polygons, with the spatial features stored in the column “the_geom,” the following SQL statement returns the area of the city of Seattle in hectares:

```
SELECT ST_Area(the_geom)/10000 AS hectares
FROM Cities WHERE name = 'SEATTLE'
```

Using a spatially enabled table containing river features, a slightly more complicated query returns the length of river segments in each state they pass through:

```
SELECT r.name AS river, s.name AS state,
ST_Length(ST_Intersection(r.the_geom, s.the_geom)) AS length
FROM Rivers r, States s WHERE
ST_Crosses(r.the_geom, s.the_geom) = 'T'
```

Spatial databases can perform many additional operations, such as coordinate transformations, buffering, and overlays. The results of these operations can be output as plain text or written into a new table to be used in further spatial analyses.

Dylan Keon

See also Data Querying in GIS; Enterprise GIS; GIS Design; GIS Software; Spatial Data Models; Spatial Data Structures
DATABASE VERSIONING

A spatial information system provides storage and management for data such that it has integrity in the representation of information. Integrity refers to the fact that the information stored in the system represents accurately the physical world. After all, a computer’s data are an abstraction of reality. To develop integrity, an information system must have a data model describing the relationships among data and the types of data being stored. This approach has come to define, in conjunction with the actual data being stored, what an information system is. The information system also needs to support several practical considerations from a user’s point of view and an administrator’s point of view. One of these is the need for versioning and annotation of groups of related information, such as multiple spatial data sets about the same subject. For example, multiple data sets about town X can exist covering a period of time.

The design of versioning and lineage into a spatial information system can be accomplished in its underlying data model via a recursive relationship on an entity. This entity describes groups or collections of related, but disparate, spatial information. Such information might include vector and raster data, aerial photographs, and ground observations. For discussion purposes, let us call this entity Set.

Typically, versioning of groups of spatial information occurs as descendant versions of a set are created. As versions of a set are created, the relationship of the sets to each other can be described by a lineage tree. The versioning lineage tree is composed of ancestors and children derived from a single ancestral spatial data set entity instance. All children in the lineage tree and thus the sets described by the tree represent collections of information about a spatial extent of data at a given point in time.

The physical implementation of the set concept and lineage trees could be created by putting all data for a spatial set into a given directory on the computer system. A collection of these directories represents a lineage tree and versions of the same spatial data set. These directories could have a flat structure or a hierarchical structure.

Hierarchical structure is the most natural method and closest in structure to the concept of lineages where children folders are descendants of the root spatial set. However, as children are versioned, this physical implementation will rapidly lead to a directory tree that can be very deep and unwieldy. The concept of the set entity in the data model and its physical implementation is to contain descriptive information about all the types of data stored in a given directory. The types of information that could be described and stored should at a minimum include the data files, schema files, SQL procedural code, annotational information about use, version number for the data, security policy, backups, and the types of users with access of use to the set data.

At some point in the life cycle of a set, a user may create a new branch of the lineage tree that may represent a different version of the original data. With multiple branches for a given collection of spatial information, versioning management systems should allow a branch to be merged back into a parallel branch that it was originally derived from. Merging branches can be done easily if there is no spatial extent overlap in the data sets. If there is overlap, it still can easily be done if the objects in the versioned sets have not been edited in such a way that an object in set A did not exist in set B when the spatial set was originally versioned. Of course, merging of sets with no common ancestor should be relatively straightforward.

In the set management approach a set should be considered and treated as a single spatial
information entity, but it is in reality a collection of individual data files, each of which may be changed to make the entire set a candidate for a new version. Each set can thus be thought of as uniquely containing information denoted by a spatial extent. In a simple spatial set lineage tree with no branching allowed, it is a relatively easy problem to create new versions of a set. A user would simply create a new version and make the lineage ancestry association using the recursive link provided in the Set entity in the data model describing sets of spatial information.

Branching and versioning in the set approach is also equally straightforward. Versioning, the creation of a new branch, could occur for a variety of reasons but would most likely occur to express a change in the user’s ownership of a given set. In this case, new versions can only be created from the tip of a lineage tree’s chain (i.e., the most recent set of data files). If a new user wanted to work on a set further back in the lineage set tree, a new branch could be created at that point, which then becomes a parallel lineage chain as shown in Figure 1.

In Figure 1, the centerline is the primary lineage for a set; each circle represents a new version of the original set. Versioning can only occur at the tip of each branch; however, parallel lineages can be created of the set’s information. Each node in Figure 1 represents a set of related spatial information that can be implemented in the physical directory structure mentioned previously. Note that because the versions of set data are derived from the same original set, merging of the lineage trees needs to be done carefully to prevent corruption of the integrity of information as mentioned in the previous discussion. For example, spatial object A in the main lineage and its copy in A’ could be potentials for information corruption if both of them are modified with different information and then the lineages are merged. If the information in A and A’ are in conflict, the question of which contains the correct information arises. This could be resolved by setting a rule that when there are conflicts in spatial object information, the object on the main lineage is accepted as being the accurate information. Of course, other priorities could be established for management of merge conflicts.

The main goals for the application of the set concept in versioning and version management are to

- allow for the existence of different ownerships of information,
- maintain integrity in information, and
- manage and point directions to how merging of different lineage version might be accomplished for various types of spatial set relations.

In summary, the application of the set concept, versioning, and the lineage tree can be a powerful paradigm for the management of multiple collections of spatial information. It allows aggregation of disparate types of spatial information into a set model that can be managed, suggests an implementation structure, and allows different versions of sets containing spatial information to coexist peacefully without the loss of data integrity.

Gregory L. Vert

See also Business Models for Geographic Information Systems; Client-Server Architecture; Critical GIS; Database Management Systems; Geographic Information Systems; GIScience

Further Readings


**DATA CLASSIFICATION SCHEMES**

In the graphic display of quantitative information, various classification schemes are used to organize data for meaningful interpretation. The human mind makes sense of the world by grouping and ordering what the senses perceive. The mind *groups* elements such as plants, animals, clothing, and so on in processes of qualitative judgment. It *orders* elements such as big trees, small trees, and so on in processes of quantitative judgment. Grouping and ordering are fundamental to classification schemes used to prepare data for mapping.

A set of data is a selected group of elements—a “population.” This population is first ordered (arranged along the number line from smallest value to largest) and then divided into classes of closely related numeric value. Each class is then represented on the map by a distinctive symbol. The process of assigning data elements to a class can be accomplished with simple arithmetic processes or with statistical processes.

**Figure 1** Classification by natural breaks

*Source:* Cartography by Alex Feldman; data by U.S. Census American FactFinder.

**Figure 2** Classification by equal interval

*Source:* Cartography by Alex Feldman; data by U.S. Census American FactFinder.
A simple, yet effective, arithmetic process is the division of the data by *natural breaks*. Once arranged along the number line, similar-value data elements are assigned a class. Each class is then assigned a specific color symbol and mapped. Using this classification scheme, a simple choropleth (value by place) map results (Figure 1).

The data can also be classed by *equal interval* (Figure 2). In this method, the data are divided into an arbitrary number of classes (percentiles, deciles, quintiles, etc.), and the same number data elements are grouped into each class. The arbitrary assignment into equal-interval classes, however, frequently obscures significant differences in data values. Compare Figures 1 and 2.

Another simple, yet effective, classification method is the *interquartile range*. In this method, the data are grouped into four equal intervals (quartiles). The middle intervals are collapsed into one class—the interquartile range. This interquartile range is assigned one symbol, while the first and last quartiles are assigned other symbols. The resulting map is a three-class map that emphasizes the lowest and highest values of the data set. The effectiveness of this method is seen in Figure 3, where the richest and poorest areas of Tennessee are easily identified.

For a more nuanced classification, the *standard deviation* is determined for each data element, and then groups of similar standard deviation are clustered into classes. While this is a preferred method for many mapping tasks, a thorough interpretation of the map is not possible for the reader with little knowledge of descriptive statistics. A map classed by standard deviation can still

**Figure 3** Classification by interquartile range

*Source: Cartography by Alex Feldman; data by U.S. Census American FactFinder.*

**Figure 4** Classification by standard deviation

*Source: Cartography by Alex Feldman; data by U.S. Census American FactFinder.*
be a powerful communication tool if it is properly symbolized. As Figure 4 demonstrates, the areas of greatest and least poverty are still discernable to the reader aware that color differences represent value differences.

Other and more complex classification schemes can be found at various sources.

Elizabeth Vaughan

See also Choropleth Maps; Map Design; Map Visualization

Further Readings


DATA COMPRESSION METHODS

Data compression is a common solution to two problems: (1) the need to decrease the storage space required for spatial data and (2) the need to minimize the resulting transmission times. Extensive exploratory research has been carried out and many algorithms have been implemented for the compression of image data and video data. Several industrial standards, such as JPEG2000 and MPEG4, have been approved as well. Three factors have to be carefully considered when the user selects or implements one data compression method, namely, the degree of compression, the amount of “lossy” information, and the time required to compress and decompress the data.

Data compression can be considered to be “lossy” or “lossless.” Lossless compression methods (e.g., Huffman coding, arithmetic coding) first detect the probability of each symbol and then encode them according to some coding algorithms and require a decoding algorithm to reconstruct the original data from the compressed data. The reconstructed data are identical to the original.

Several solutions can be directly applied for lossless compression, such as running length encoding, dictionary coders (e.g., LZW), the Burrows-Wheeler transform, context mixing, and Slepian-Wolf coding. For example, Huffman coding-based algorithms create a Huffman tree according to the frequency or probability of a symbol’s appearance and encode high-probability symbols with low bits. Arithmetic coding-based algorithms represent each possible sequence of \( n \) symbols by a separate interval on the number line between 0 and 1. Compared with lossless compression, lossy compression is another kind of compression method that is possible if some loss of fidelity is acceptable. Lossy compression methods include the discrete cosine transform, fractal compression, wavelet compression, vector quantization, linear predictive coding, Modulo-N code for correlated data, A-law Compander, Mu-law Compander, and Wyner-Ziv coding.

Generally, a lossy data compression will be guided by research on how people perceive the data in question. Compression methods are commonly used in conjunction with two representations of spatial data: raster data and vector data. Various sophisticated algorithms have been implemented for the compression of raster data (e.g., remote sensing imagery, digital elevation models), such as wavelet-based compression algorithms, and for the compression of triangulated irregular network (TIN)-based models (e.g., terrains). Nevertheless, algorithms for compressing raster data and TIN-based models cannot directly compress vector data because of the intrinsic complexity of topology within vector data. As a consistent topology has to be maintained, viable solutions for the compression of vector data should be carefully considered. Alternatively, the time performance of decompression is also an important factor.

Vector quantization (VQ) is a lossy data compression method based on the principle of block coding. It is a fixed-to-fixed length algorithm. The main principle of VQ is as follows: Given a vector source with its known statistical properties, a distortion measurement and the number of vectors find a dictionary (codebook) that results in the smallest average distortion. Therefore, the procedure of VQ is to create a dictionary with \( S' \) symbols from the message with \( S \) symbols \( (S' < S) \) and to build a mapping between \( S' \) and \( S \). VQ-based algorithms are more suitable for the compression of points in geographic data layers.
Besides the degree of compression, the amount of lossy information, and the time performance of compressing and decompressing data, the consistency of topology and the data quality of decompressed data must be considered for selecting or implementing a viable compression solution. The decompressed data should maintain the maximum similarity of geometric shapes and maintain a consistent topology; otherwise, the compression method will have difficulties in meeting user requirements even if it enables a high compression ratio.

A key factor for vector data compression is to find a subset of the original vertices to represent a polygonal curve, and maintain an optimal approximation of the polygonal curve and topological consistency of a set of polygonal curves (e.g., vector map data). Bisheng Yang, Ross Purves, and Robert Weibel recently proposed a variable resolution compression method for vector maps. The main idea of the proposed method is to first transform the compression of the lines constituting polygons to vertices. Then, a VQ method is adopted to compress the vertices. The proposed method

- preserves consistent topology within the compressed results for a correctly selected distance threshold,
- compresses vector data at different compression ratios according to varied requirements, and
- achieves a reasonable balance between the compression ratio and the distortions of geometric shapes.

Data compression methods, particularly for the compression of vector data, are receiving more attention because of increasing requirements in Web-based GIS applications and geospatial information services. The compression of vector data is central to mobile visualization, Internet transmission, and interactive visualization due to increasing data volumes, narrow network bandwidths, and limited screen sizes. One of the key challenges for current and future research efforts is fast compression and decompression of vector data online to facilitate real-time mapping and visualization of geospatial data.

Bisheng Yang

See also Data Editing; Digital Terrain Model; Geospatial Industry; Internet GIS; Triangulated Irregular Network (TIN) Data Model; Vectorization; Web Service Architectures for GIS

Further Readings


Simply stated, *data editing* refers to the activity of detecting and rectifying errors in data. This could also involve updating a data set to make it more contemporary. From a geographical perspective, data collected could be spatial (locational) in nature, or they could be aspatial (descriptive) data. For example, the coordinates (e.g., latitude/longitude) of groundwater monitoring wells are spatial data, while descriptive data about groundwater depth, yield, and so on would be aspatial data.

There are several ways in which geographical data collected can be in error or in need of updating. Errors may arise from a systematic shift in measurements due to calibration procedures, they may be random in nature, or they may be due to errors in data entry (coding) or updating methods. Usually, systematic errors manifest as a set of values either above or below an expected value, while random errors can be seen as outliers. However, there are times when an outlier could be real and indicate a dip or a spike in the measured value, for example, a soil sample taken over a sweet spot indicating a high mineral ore concentration.

Correcting errors in data involves evaluating the magnitude of error by comparing the measured value with a reference value. This task is known as accuracy assessment. In general, reference data constitute values collected at a higher degree of accuracy and precision. For example, locational
Data collected using a GPS unit with a higher precision (decimal degrees) and accuracy (submeter) could be used for accuracy assessment. Additionally, high-resolution aerial images (georeferenced) could be used in assessing the accuracy of location data. On the other hand, aspatial data could also be collected at higher levels of precision and thus serve as references, for example, soil samples analyzed for plant nutrients using high-precision instruments. Additionally, the error might be due to data duplication (recorded more than once), inconsistency (conflicting values), deviation from historical trends, and so on. Thus, error detection is basically applying “rules” to data, and these rules are usually formulated by expert knowledge.

Rectifying systematic errors in spatial data involves adding or subtracting the delta value (error between the reference and measured values) from the location data. This is similar to systematically shifting the longitude ($x$), and latitude ($y$) data coordinates. Editing aspatial data ranges from manual editing of the data values to writing simple programs (macros) that seek and fix the errors present in attribute data. Also, this process could involve a statistical edit wherein the attribute value in one record is compared against a frequency distribution for that attribute for all other records. So how exactly do we fix the erroneous value? Well, we could delete the erroneous value and thus eliminate it from our subsequent data analysis. Alternatively, imputation procedures could be used to fill in gaps or missing values. We could interpolate a simple mean for the erroneous or missing value based on its immediate neighboring values or class. Advanced interpolation techniques take into account the magnitude and direction of the neighboring values in the interpolation process. Also, specialized programs (expert systems) can automatically impute data based on a knowledge base or a statistical method.

Thus, data editing is a valuable data management operation that ensures data quality by providing data that are accurate, consistent, and complete. This greatly enhances the validity and credibility of subsequent analyses implemented using the data.

Mahesh Rao

See also Analytical Operations in GIS; Database Management Systems

Further Readings


Data format conversion refers to the conversion of digital data from one encoding scheme to another. Common examples include the conversion of a Microsoft Word document to Adobe PDF format, a TIFF image to JPEG format, and a WAV audio file to MP3 format. Many popular software packages provide Save As, Import, and Export functions, all of which perform data format conversions. While certain data conversion operations retain all data (lossless conversion), others result in an irrecoverable loss of data (lossy conversion). For example, when an uncompressed image (e.g., TIFF format) is converted to JPEG (a lossy format), certain bits are eliminated during the conversion, resulting in a permanent loss of data and a concomitant reduction in file size. Some lossless conversions (e.g., of a WAV audio file to FLAC format) use data compression techniques, achieving smaller file sizes while retaining information that allows full reconstruction of the original file.

Data format conversions are common operations in geographic information systems (GIS). The conversion of a delimited text file containing point data (coordinates and attributes) to the shapefile format is an example of a lossless conversion—no data are lost during the conversion. Likewise, data collected by a GPS unit can be converted to a shapefile or other vector format with no loss of information. A common GIS operation is the conversion of vector data to a raster format.
DATA INDEXING

DATA INDEXING

for use in map algebra and other raster analysis techniques. The conversion of vector data to a raster representation is often a lossy operation; for example, data would be lost if a line layer of roads data collected using a commercial-grade, submeter-accuracy GPS was converted to a raster data set of 10-meter resolution. Vector data can also be converted to other vector formats. For example, the conversion of an ESRI coverage to a shapefile preserves feature and attribute data, but results in a loss of topological information. Information can also be lost when a geographic data set is converted to another representation of lower resolution and less information content, which often results from a change in scale.

Most GIS software programs can perform common GIS data format conversions. The open source Geospatial Data Abstraction Library is capable of converting nearly 100 raster data formats to other supported raster formats, and its analog, the open-source OGR Simple Features Library, can convert many vector data formats to other supported vector formats. Other specialized GIS data conversion tools exist, such as shapelib, a library for manipulating shapefile data sets. Transformational languages such as variants of the Extensible Markup Language (XML) can also be used for data conversion. One example, the Geographic Markup Language (GML), is a common format for transformation and exchange of spatial data sets. Tools such as OGR enable the conversion of shapefiles and other vector data formats to GML, which can then be manipulated or directly converted to another vector file format or spatial database. Data format conversions can also be accomplished within databases, using standard functions such as to_number(), which converts a text string to a numerical representation.

Dylan Keon

See also Data Compression Methods; Data Editing; Spatial Data Structures

Further Readings


DATA INDEXING

Stripped of its numerous complexities, data indexing is the high-tech version of the tedious indexing process familiar to any 20th-century textbook author and results in the same benefits for the reader. The major difference is that computers are doing both the indexing and the accessing of the indexes. Just as with a textbook index, the basic goal behind data indexing is to prevent the reader (in this case, the computer) from having to search and sift through an entire data set—a prohibitively slow process—to find only the specific information it needs. More formally, this allows the reader random access to indexed elements, which is much more efficient than sequential access. Data indexes, crucially important data structures in computer science, make possible a vast number of technologies, particularly those dealing with the enormous information repositories so prevalent in the modern age. Without them, everything from a Google search to the dynamic display of geographic information would be exponentially slower, easily to the point of total impracticality. Two types of data indexing are prevalent in geography. The first is “nonspatial indexing,” and the second is “spatial indexing.” Nonspatial indexing is applied throughout computer science and related fields, while spatial indexing is designed specifically to work with n-dimensional data, where n ≥ 2 (in geography, we usually find that n = 2 or 3).

Nonspatial Data Indexing

As noted above, nonspatial indexing appears in the implementation of a very large percentage of computer applications. While there are many types of nonspatial indexes, they all allow a program to query a data set or database for specific information much faster than would be possible if the computer had to look through the entire data set or database for that information. The type of index included in a particular application is usually carefully chosen
based on certain key properties, such as the additional data storage costs (just like a book’s index, data indexes take up additional space), the type of data that will be indexed, and, of course, the speed of lookups. Sometimes the indexing process and the data index type can be very simple. For example, the file with the “.shx” extension that is part of the shapefile specification is purely a rudimentary index file that makes it easier for any program reading the shapefile to find the data related to a specific point, line, or polygon. It is essentially made up of a set of pointers to the beginning of data records in the .shp file.

Many times, however, more complex data index approaches are needed. One important type of nontrivial data index is the B-tree index. Invented in 1972, B-trees and their variants are still vital in modern systems. For example, B-trees are the default nonspatial indexing method in the very popular MySQL database management system. While the details of B-tree implementations are outside the scope of this text, the basic idea of their structure is illustrated in Figure 1.

The shaded entries indicate which “nodes” need to be examined to find the record number “18” (and data about record 18). A “sequential scan” of the “data set” would have had to look at 18 different entries (assuming the data had been presorted), while the B-tree reduces that number to 3.

Spatial Data Indexing

Spatial data indexing was also in use long before the computer became a primary method of storing and analyzing spatial information. In fact, the indexes at the back of atlases, with page numbers and grid cell identifiers given for each entry, are an (inverted) elementary form of a spatial data indexing methodology called “grid indexing.”

Currently, spatial data indexing is frequently implemented through an R-tree data structure or a derivative. MySQL, for example, uses an R-tree implementation for indexing spatial data, and Oracle Spatial, IBM Informix, and PostGIS implement them as well. R-trees, originally presented in 1984 by Antonin Guttman, effectively allow the efficient answer of queries along the lines of “Retrieve all school objects within 2 miles of my current location.” Instead of searching through all entries in a database of school objects, which may number in the millions, the implementing application can access the R-tree index and can usually receive the answer in exponentially fewer steps along the same lines as seen in the B-tree example.
A popular variant of the R-tree is the R*-tree. Themselves a variation on the theme of the B-tree, R-trees essentially split a given space into nested or overlapping minimum bounding rectangles (MBRs), known to geographers as bounding boxes, and store data entries hierarchically based on these MBRs.

While R-trees (and derivatives) are likely the most common spatial data indexing methodologies in the computer science world, other methodologies are optimal under certain conditions that appear frequently in geography contexts. The grid index, an elaboration on the aforementioned “low-tech” atlas index methodology, is easy to understand and create, is relatively fast, and can support many types and densities of data. As such, grid indexes are quite often found in parts of commercial GIS, including many areas of ArcGIS (as of Version 9.3). Quadtrees, specialized forms of the grid index and well-known index types in and of themselves, are used by GIS developers when they need to consider many types of geographic data—particularly both vector and raster data—and need do so in a relatively easy-to-implement and efficient fashion.

Brent Hecht

See also Database Management Systems; GIScience; High-Performance Computing; Spatial Data Infrastructures

Further Readings


A database system is an essential component of a geographic information system (GIS). GIS are driven off databases. Almost all entities in the world are tied to some geographical location, expressed by x (longitude) and y (latitude) coordinates. These entities are stored in GIS as a series of points, lines, and polygons.

The power of a GIS is its ability to store the coordinate information; however, most GIS have the ability to connect to other information stored in a database management system (DBMS). For example, usually one land parcel within a city is more than likely owned by one person; however, if that land parcel is located in Manhattan, with the dollar value of it being 20 million, more often than not that land parcel is likely owned by more than just one person. We can store information about this land parcel, such as acreage, asset value, and ownership information, in a single database table, but some of the information would be duplicated. To help prevent this duplication, a good database management practice is to create a unique key that distinguishes each land parcel from other parcels. To build a database dealing with land parcels, a good practice is to assign a parcel identification number to each land parcel. This step ensures that the land parcel can be distinguished from all other land parcels.

Instead of storing all the information in one table, where duplication can be a problem, relational database techniques can be applied and tables broken up into two or more tables. Certain unique fields within the tables are assigned a key that uniquely identifies them. Once unique key fields are established in the tables are assigned a key that uniquely identifies them. Once unique key fields are established in the database tables, most DBMS have functions called joins and links that allow many tables to act as one table, thus preventing duplication of data. Most GIS have join and link functions built within the software. For example, in most GIS one can select a single land parcel and link it to the ownership table and return all the current owners for that land parcel.
Relational database standards have facilitated querying and retrieving data from database tables. In particular, one such standard called the SQL language has been one of the major reasons for the success of relational databases. So what is SQL? The acronym SQL stands for structured query language. Originally SQL was called SEQUEL (for structured English query language). Over the years several versions of the SQL standards have evolved, and several database vendors have created their own version of SQL or optioned a current standard.

The power of SQL is its ability to extract information from the database based on some user-defined criteria. For example, if you wanted to get all land parcels with acre size greater than 5, you would create an SQL query such as “Acres” > 5. In the early days, before today’s easy-to-use Application Programming Interfaces were created to construct SQL queries, querying databases was time-consuming and challenging. Today, several advanced SQL tools are in use that have a user-friendly interface. Figure 1 is an illustration of a database query interface created by the Environmental Systems Research Institute, the leader in GIS technology.

What makes GIS technology stand out from other database systems is its ability to perform a spatial query. A spatial query is a special type of database query supported by geodatabases. Spatial queries differ from SQL queries in several important ways. Two of the most important are that they allow the use of geometry data types such as points, lines, and polygons and that these queries consider spatial relationships between the objects of query.

For example, you work for a highway district as a planner and a new proposed highway plan passes through a number of land parcels. You want to get in touch with the owners of these land parcels to coordinate preparations for the proposed highway plan. To do this using a GIS, you could perform a query, in which a spatial relationship between the new highway map and the land parcels would be represented by a line-in-polygon intersection. The result of the query would include the contact information for those land owners whose parcels are to be intersected by the proposed highway.

Tyson Taylor

See also Analytical Operations in GIS

DATUMS

A geodetic datum is a fixed reference point according to which the size and shape of Earth are measured. They are also used in measuring gravitational pull and rotation. There are at least three kinds of geodetic datums used in geography. Vertical datums form the zero surface for vertical measurements of altitude and elevation. Horizontal datums define the size and shape of Earth and the origin and orientation of a horizontal coordinate system. Complete datums can provide both vertical and horizontal zero origins and coordinate system definitions. Some complete geodetic datums define Earth’s shape, horizontal and vertical coordinate system origins, gravity fields, and physical constants such as rotational velocity and Earth’s gravitational constant.
Without a vertical datum, the altitude of a point is ambiguous, and without a horizontal datum the longitude and latitude of a point are not sufficiently well-known to locate it on the Earth. With the specification of a vertical zero, such as the ground level, the mean sea level, or the surface of a particular reference ellipsoid, the elevation of a point can be known. If the equator is defined as the origin for lines of latitude, and Greenwich, England, is defined as the origin for lines of longitude, then any point can be found to at least within a few kilometers on the globe. To specify a position to within a few meters, coordinates must be associated with specific geodetic horizontal and vertical datums.

Since the French geodetic expeditions of the mid 1700s to Peru and Lapland, Earth has been considered to be a slightly flattened sphere, a spheroid, an ellipse of rotation, or an ellipsoid. There have been dozens of different Earth shapes on which mapping systems and coordinate systems have been based. While there have been other Earth shapes proposed and used in geodesy, most are specified as ellipsoids with an equatorial radius and either a polar radius or a parameter that represents the relationship between the equatorial radius and the polar radius.

The equator forms the zero point for lines of latitude. There is no equivalent physical feature on Earth that naturally lends itself to the origin for lines of longitude. The zero point, or zero line of longitude, the prime meridian, has been defined at different times by different interests to pass through 1 of more than 20 cities, including Moscow, Paris, Madrid, Rio de Janeiro, Lisbon, the Canary Islands, Washington, Tokyo, and Greenwich, England. A geodetic datum defines at least a single monument with a specific longitude and latitude as well as the distance and direction to another monument.

Most modern datums use a network of monuments to define a system of longitude and latitude or vertical heights over regions covering small islands, countries, continents, or the entire Earth. The Cape Canaveral Datum is a local datum defined for use in Florida and the Bahamas. The North American Datum of 1983 (NAD83) is a regional horizontal datum. The U.S. National Geodetic Vertical Datum (NAVD 88) is a regional vertical datum. Geodetic Reference System 1980 (GRG80) is a global horizontal geodetic datum. World Geodetic System 1984 (WGS-84) is a complete global geodetic datum with both horizontal and vertical components, as well as Earth’s gravitational and rotation parameters.

Peter H. Dana

See also Altitude; Earth’s Coordinate Grid; Geodesy; Latitude; Longitude

Further Readings


William Morris Davis is the single most influential geomorphologist to have written in the English language, and perhaps in any language. Given the enormous growth and diversity in the discipline, it is unlikely that anyone will ever again dominate it to the degree that he did. The primary component of his reputation is the creation and indefatigable proselytism of the geomorphic cycle, also known as the geographical cycle. The model attempts to provide a comprehensive theoretical explanation of the development of landforms and many of their individual elements during the perceived lengthy periods between regional uplifts. The model gained enormous, but not unchallenged, currency and represented the prevailing orthodoxy in the English world from about 1900 until around World War II. It was challenged in research circles starting just before the war and more vigorously
DEAR, MICHAEL

thereafter; in some places it lingered longer, and in the realm of teaching, considerably longer.

The model itself was supported by Davis’s prolific publication record and introduction of innumerable technical terms. Many such terms are morphogenetic; that is, they not only describe how something looks but also purport to explain its origins, peneplain being an exemplar.

Davis was reared in an influential East Coast Quaker family, attended Harvard University, and was on the faculty from the late 1870s until 1912. He conducted fieldwork on every continent save Antarctica. His numerous professional honors included being president of both the Association of American Geographers and the Geological Society of America. His role as a bridge builder between the two disciplines is an important part of his personal story.

Davis was a prodigious worker, writing hundreds of articles, introducing well over 100 technical terms, accepting numerous speaking engagements on both the national and international levels, and conducting what today would probably be called outreach (i.e., talking to the nonscientific community). An additional personal attribute was his real flair for sketching. His beautifully conceived and executed diagrams essentially present his geomorphological ideas in a separate language. His influence in retirement was considerable because he moved to the Western United States, where he inspired many students of geology in his post-Harvard years and exhibited much greater intellectual flexibility than is often recognized.

Whatever limitations are now assigned to Davis and his geomorphic cycle, it is appropriate to view the latter as a monumental intellectual achievement: influential when it dominated and also a powerful intellectual challenge to be satisfactorily met for those who followed and sought to modify the framework of geomorphology. Furthermore, Davis provided a powerful personal aura of scientific professionalism in an embryonic discipline.

Colin Edward Thorn

See also Association of American Geographers; Geomorphic Cycle; Geomorphology; Gilbert, Grove Karl; Penck, Walter

Further Readings


Michael Dear has been a significant force in human geography and urban analysis for close to four decades. A provocative and prolific writer, Dear has authored or edited 19 books and more than 100 scholarly articles, covering topics, including the mentally ill, homelessness, the state, social theory, postmodernism, and, most recently, comparative urbanism.

Born in Treorchy, in the Rhondda Valley of South Wales, Dear obtained a BA in geography from the University of Birmingham (1966), an MPhil in town planning from the University of London (1966), and MA (1972) and PhD (1974) degrees in regional science from the University of Pennsylvania. He taught at McMaster University in Hamilton, Ontario, Canada. He was a professor of geography and urban planning at University of Southern California for many years before moving to the department of urban and regional planning at the University of California, Berkeley in 2009. He has held fellowships at the Center for Advanced Study in the Behavioral Sciences at Stanford University and the Rockefeller Center in Bellagio, Italy, and was a recipient of a Guggenheim Fellowship.

Dear has remained steadfast in his attempt to refashion geography to align more closely with social theory. Throughout his career, he has been a strong advocate and outspoken proponent in the development and promotion of key theories and methods, including Marxism, structuration, and postmodernism. With respect to postmodernism,
Dear argued that its emphasis on difference places an intense spotlight on the subject, the body, and personal identity. Consequently, postmodernism is a powerful means to examine questions of social justice. A concern with policy issues remains a strong focus in his entire oeuvre.

Dear is the founding editor of the journal Environment and Planning D: Society and Space (with Edward Soja, Allen Scott, Michael Storper, and Mike Davis) and is a leading exponent of the “Los Angeles School” of urbanism. Interdisciplinary in his work, Dear draws insights from and has contributed to geography, planning, social work, psychiatry, public health, history, and political science.

Not averse to controversy, Dear has been a highly visible and vocal participant in the ongoing debates and discussions shaping human geography throughout the late 20th and early 21st centuries. In 1988, Dear issued a “postmodern challenge” to human geography. In his “reconstructed” human geography, the objective should be to “understand the simultaneity of time and space in structuring social processes” (p. 270). More recently, Dear has called attention to the “politics of hate” that he argues may be found in the discipline.

James Tyner

See also Los Angeles School; Postmodernism

History

In the 1970s, oil-producing countries increased the price of oil significantly, generating massive profits. These profits, referred to as petrodollars, were deposited in a number of Western banks, especially in the United Kingdom and the United States, and they were recycled by bankers in the form of loans to developing countries. Many of the loans were channeled to Latin American countries as well as to several countries in Asia and sub-Saharan Africa. In some cases, money was lent to military dictatorships that had never been democratically elected. They used the loans in wasteful and unproductive ways, including the creation of large steel plants that never produced steel, and environmentally damaging dams, and to purchase weapons. At the time, however, interest rates were low and commodity prices were rising. External debts, while large, could be serviced with export revenues.

This situation changed in the 1980s, when conservative world leaders such as Ronald Reagan and Margaret Thatcher began to put restrictive monetarist policies in place, which led to increases in interest rates. Commodity prices began to fall, while the prices of manufactured goods began to rise, which resulted in deteriorating terms of trade for commodity-producing developing countries. As a result, many borrowing countries found it increasingly difficult to meet their debt service obligations.

The worsening situation came to a head in August, 1982, when the Mexican government announced that it had run out of foreign exchange reserves and would be unable to service its debt. The bankers reacted promptly to the Mexican default, fearing its repeat across the continent.
The 1982 debt crisis created a new role for the International Monetary Fund (IMF), which was charged with managing the crisis. The banks were forced by the IMF to contribute to the Mexican bailout and that of other countries that were running into difficulties. Lengthy negotiations were held with countries on a bilateral basis to prevent the formation of a debtors’ cartel and to save the banks that had lent the money.

The response to the question of default was to provide fresh loans so that debtor countries could continue to service old loans and to keep the dollars flowing from the developing countries to the northern banks. Effectively, the debt crisis was “resolved” by further increasing the already unpayable debts of Third World countries. These actions converted the poorer countries of the South into net exporters of capital to the wealthier countries of the North. By 2000, many countries of the South had much higher total debt burdens compared with two decades earlier.

Although the debt crisis sent the banking world into turmoil, for some economists it was seen as a blessing in disguise. At this historical moment, neoliberal thought, a belief in the economic efficiency of the market and the need to reduce the role of the state in economic affairs, was once again in ascendency and was being promoted by influential economists such as those of the Chicago School of Economics. The debt crisis was seized on as a means to reorient Third World economies toward more promarket economic policies and dismantle protectionisms and state subsidies. Bailout packages were thus accompanied by strict conditionalities in the form of structural adjustment policies, and Third World countries were discursively constructed as sites of economic inefficiency and malaise that required harsh economic medicine.

Structural Adjustment

After the debt crisis of the 1980s, structural adjustment became the dominant development model, promoted and enforced by the institutions of global governance. In return for new loans, countries are required to adopt a series of economic measures that involve trade liberalization, an emphasis on export production, and the withdrawal of state intervention. Import tariffs are lowered and economies are opened up to transnational capital. The local currency is devalued to stimulate exports. State-owned enterprises are privatized. State subsidies on food or fuel are removed, and cutbacks in public spending are made. Essentially, countries are required to earn more and spend less to keep debt service payments flowing.

The social costs of the debt crisis are well documented. There is no doubt that the burden of debt has been borne by the poorest and most vulnerable sectors of the population. In the 1980s and 1990s, several countries were spending more of their export revenue on debt service than on health care and education. In many countries, rapid trade liberalization resulted in job losses and an increase in poverty and malnutrition. Working people saw their real wages and purchasing power plummet. The emphasis on exports devastated small farmers who produce food for domestic consumption. These farmers found themselves deprived of access to credit, and many have been displaced from their land by large export-oriented agribusinesses. The export-oriented economy has also hastened environmental degradation in many parts of the world, given its dependence on resource extraction and monocultures that require an intensified use of chemical fertilizers and pesticides. Women, too, have been particularly negatively affected as the cutbacks in health care and education and the removal of subsidies on basic foodstuffs complicate the task of feeding families and caring for children and the elderly. The human cost of the debt crisis was so severe in Latin America that the 1980s are often referred to as the “lost decade.”

Responses and Solutions to the Crisis

A number of official solutions to the debt crisis have been proposed by First World governments and the international financial institutions. The 1985 Baker Plan was focused on further lending with strict conditions attached, and so it further institutionalized the implementation of structural adjustment. The 1989 Brady Plan, also conditional on economic reforms, achieved limited debt relief through a range of mechanisms that converted debt into tradable bonds. While these two plans restructured and diversified the debt to
DEBT AND DEBT CRISIS

some extent, they failed to achieve sustainable debt relief.

In 1996, the IMF and the World Bank introduced the Heavily Indebted Poor Countries (HIPC) initiative as a means to provide debt relief for some of the world's poorest and most debt-laden countries. The vast majority of HIPC countries are in sub-Saharan Africa. In Latin America, Honduras, Nicaragua, and Bolivia also qualified for the HIPC.

To qualify for debt relief, countries are required to produce a Poverty Reduction Strategy Paper (PRSP) in consultation with civil society. The savings in debt servicing are then to be channeled into poverty reduction. While theoretically providing much needed debt cancellation of up to 80% of a country's debt, the HIPC initiative has attracted significant criticism from NGOs, lobby groups, and intellectuals. Despite the shift in language, the HIPC continues to place emphasis on macroeconomic stability and growth. To access the HIPC, countries must continue to adhere to strict adjustment and austerity measures, and countries that do not meet these conditions and fail to satisfy the IMF can be and sometimes are refused debt relief. Criticisms have also been leveled at the length of time it takes for a country to meet the completion point (at which debt is cancelled), the fact that only a part of a country's external debt is cancelled, and the limitations and weaknesses in the model of consultation. Critics have also called into question the IMF and World Bank definition of debt sustainability, which means that many countries with considerable debt burdens are unable to enter the HIPC.

Nevertheless, the HIPC does at the very least recognize that debt cancellation is the only viable and long-term solution to the problem. Over the past decade, some lender countries have also canceled their bilateral debts with some of the most indebted states. Full cancellation of external debt continues, however, to be resisted by the multilateral institutions as it would remove the leverage that these institutions exert over domestic economic policy in the Third World.

The latest official approach to debt relief is the 2006 Multilateral Debt Reduction Initiative (MDRI), developed by the G8. The MDRI strengthens the focus on the cancellation of multilateral debt but has a number of limitations. It continues to insist on the completion of HIPC conditions, it only includes a small number of countries, and it only includes debts to the IMF, World Bank, and the African Development Fund. It does not include bilateral debts that have not been cancelled, commercial debts to private banks, IMF debts incurred after 2003, World Bank debts incurred after 2004, and debts to other multilateral lenders such as the Inter-American Development Bank. Debt relief has also been accompanied by a reduction in aid allocations.

Despite the flaws of the existing mechanisms, some countries have begun to invest in health care, education, and infrastructure as a result of the partial reduction of their unpayable debt burdens.

Resistance to Third World Debt

There has been widespread political activism across the world as a result of Third World debt since the 1980s. In the Third World, there were many forms of resistance to neoliberal structural adjustment, including riots against the IMF and the creation of coping mechanisms such as communal kitchens where women pooled labor and resources to feed their families. In the 1980s, campaigns and boycotts against the high street banks implicated in the debt crisis were commonplace.

The economic model ushered in by the debt crisis has led to a proliferation of globalized political activism. High-profile protests in Seattle in 1999, in Cochabamba, Bolivia, in 2000, and in many other cities across the world constitute historically significant popular victories against neoliberalism and the actions of transnational capital.

The most high-profile global campaign against Third World debt has been the Jubilee Coalition, which began its campaigning in the 1990s and called for complete debt cancellation by the year 2000. Since 2000, it has continued to campaign for debt cancellation and has closely monitored the HIPC process. Similarly, the Make Poverty History campaign, organized to lobby the G8 summit in Scotland in 2005, was successful in putting Third World debt back on the political agenda.
While the PRSP process has attracted criticism, it has provided a space for political engagement in some places. In both Nicaragua and Uganda, civil society organizations have mobilized around the PRSP process to ensure savings from debt relief are channeled in a transparent and equitable way into poverty reduction programs. As a result, in some places local people have been demanding greater accountability from their national governments.

**Conclusion**

While some debt relief has been granted, campaigners continue to call for full debt cancellation as existing debt burdens continue to deprive millions of people in the Third World of social and economic opportunities. The Jubilee Campaign estimates that the world’s 50 poorest countries are still sending around US$100 million a day to rich countries in debt service and that they have a combined debt of around US$300 billion.

Neoliberal structural adjustment and financial deregulation have not prevented ongoing turmoil in financial markets. In the past decade, there have been severe financial crises in Eastern Asia, Russia, Brazil, Argentina, and Mexico. In 2008, a new financial crisis swept the world, once again brought on in part by overzealous and irresponsible lending policies of Northern banks. The 2008 financial crisis was global in its scope and had particularly harsh impacts on First World countries, including the United States and the United Kingdom. As in 1982, the official response focused on saving the banks through a financial rescue package, while ordinary people faced credit restrictions, job losses, and home foreclosures. The 2008 credit crunch, as it came to be known, is likely to undermine the ability and willingness of First World actors to find a solution to the ongoing problem of Third World debt.

Julie Cupples

*See also* Antiglobalization; Dependency Theory; Development Theory; Export-Led Development; International Monetary Fund; Neoliberalism; Structural Adjustment; Uneven Development; World Bank; World-Systems Theory

**Further Readings**


**DECOLONIZATION**

Decolonization refers to the disintegration of large empires, usually meaning the European ones that took shape starting in the 16th century, and the formal independence of former colonies. Just as colonialism began unevenly over the surface of the Earth, so too did its end. World-systems theorists argue that the opportunities for states on the global periphery to exert themselves against colonial powers are best when the core is in crisis. Thus, the Napoleonic wars of the early 19th century afforded Latin America the opportunity to break away relatively early. Similarly, World Wars I and II proved to be the pivotal moments when Western control over much of Africa and Asia was finally broken.

The shift toward decolonization during the post–World War II era was complex. Often independence movements comprised broad coalitions of nationalists, students, the intelligentsia, and peasants, frequently led by Western-educated intellectuals (e.g., Ho Chi Minh in Vietnam, Mohandas Gandhi in India). Such movements frequently had roots that extended to the 19th century and gradually grew in power over time. The Cold War rivalry between the United States and the Soviet Union afforded nationalist movements a political space that allowed them to play the superpowers off against each other as both
the United States and the USSR were eager to appear different from older European colonial conquerors and friendly to the masses of the emerging states. Often the struggle for independence was violent, involving protracted guerrilla conflicts and wars (e.g., Malaysia, Vietnam, Algeria). The relatively peaceful independence movement in India was the exception, although the country’s division into India and Pakistan involved extensive civil strife and the deaths of millions.

Independence movements succeeded in Asia and Africa throughout the late 1950s, 1960s, and 1970s, leading to a proliferation of newly independent countries (from roughly 50 in 1945 to about 196 today in the United Nations). Major milestones in this process include India and Pakistan in 1947, Indonesia in 1949, Algeria in 1964, and Angola and Mozambique in 1975. Virtually all parts of the globe today have been decolonized, with a few small exceptions (e.g., French Guiana, Martinique, Gibraltar, Tahiti, Puerto Rico).

Decolonization involved political, economic, and ideological changes. Politically, this shift brought with it a new administrative and legal apparatus in the former colony, typically modeled after the colonial one. Indeed, often the very same people who served the foreign colonial power became leaders in the newly independent one. Ideologically, decolonization opened the door to challenges to long-standing racist notions of white superiority (e.g., Ghana’s Kwame Nkrumah), allowing a variety of experimental social projects (e.g., Tanzania’s *ujamaa*, or “African socialism”).

While formal political independence inevitably brought with it the trappings of a new society—a new flag, currency, and national airline, for example—many observers pointed out that decolonization did not simply end colonial relations overnight. Indeed, to dependency and world-systems theorists, it is no accident that the former European colonies are inevitably part of the “Third World,” the enormous, diverse set of societies that encompass the vast bulk of the world’s people but relatively little of its wealth. Despite ostensible political independence, such societies were often woefully unprepared for independence economically and remained heavily dependent on their former colonial power for capital, trade assistance, and foreign aid, leading to widespread fears of neocolonialism, generally via multinational corporations. The dominant role of the United States as the world’s leading neocolonial power, in both economic and military terms, made the contrast between nominal political and substantive economic independence all the more apparent. Most former colonies have inadequate infrastructure and human capital, with economies centered on raw materials (foodstuffs, mineral ores). Regardless of the measure used—GDP per capita, energy consumption, access to health or education services—former colonies almost always lag behind the industrialized world (although some, such as Singapore, rival it, and the newly industrializing countries have made rapid progress).

Barney Warf

See also Colonialism; Dependency Theory; Neocolonialism; World-Systems Theory

Further Readings


Deep ecology emerges from the intuition that nature and self are one. It is important because it provides the motivation for much of the environmental activism that is conducted in the name of geography and education for sustainable development, and it is the source of most eco-centric (ecosystem-centered) thinking within geography. Deep ecology emerged in the early 1970s from the work of the philosopher Arne Naess. Naess sought to distinguish between environmental thinking that gave the highest priority to human welfare, which he called “shallow ecology” and that which gave priority to the welfare of the
whole ecological system, which he called “deep ecology.”

Deep ecology suggests that, even if it were desirable or ethically apt, the human species is incapable of existing in isolation. So it is in humanity’s best interest to prioritize the welfare of the life support system that it shares with all other forms of life, the living planet, ahead of any human want or desire. To aid this, John Seed and colleagues’ advice to “think like a mountain” aims to encourage the search for the viable consciousness that will help humans secure a sustainable future. Seed also promotes the connective educational exercise known as the “Council of All Beings,” where humans try to give voice to the concerns of other species.

**Ecological Self**

Deep ecology’s purpose is to make human minds fit for the biosphere. Deep ecology sees the environmental crisis as a symptom of an anthropocentric delusion that imagines the world to be merely a resource for human use. Its aim is transformational, the creation of a new worldview, through ecological self-realization, a notion that owes much to Naess’s work on the philosophies of Baruch Spinoza and especially Mahatma Gandhi’s Hindu Vedanta.

Ecological self-realization is a three-step process by which one gains self-identity with nature; it is the pedagogical framework of deep ecology. Each step involves a transformation of self-consciousness, which may be linked, metaphorically, to human maturation. Step 1 is located in childhood, when the individual self recognizes its personal autonomy and individuality, including the will to contradict. Step 2 is adolescence, when this self redefines itself in terms of a place in a social group such as family, peer group, nation, and eventually, it is hoped, the whole of humanity. In this process, the “I-self” becomes subsumed within a larger “we-self.” This is the intuition that underpins anthropocentric ecoc socialism and the doctrines of shallow ecology. Step 3, maturity, is recognition of the ecological self, where the self is redefined in terms of a role and a place within the community of all life. As Gandhi advises, all living beings are members one of another. Elsewhere, researchers discuss “egoistic” values that predispose people to protect environmental attributes that affect them personally, “altruistic” values that express concern in terms of the welfare of human society, and “biospheric” values that concern the welfare of the whole biosphere and all life, a position often criticized as antihuman or occult. Thus, ecological self-realization strives to shift the focus of the self from “ego” to “eco” and toward treating nature with the same consideration that is traditionally limited to one’s own body or immediate family. Each step diminishes the amount of the universe that is external to the s/Self. Each step also contains the realization that a human is not alone or self-sufficient and that, ultimately, humans are merely a recent addition to a much larger, respectably ancient, organic wholeness, often called “Gaia.” Here, “Earth systems science” also demonstrates that human survival depends on a myriad other creatures, both macroscopic and microscopic, some of which have long played a more important role in the global environment.

**Political Platform**

Under the influence of George Sessions, such thinking crystallized in the much criticized eight-point political platform of the deep ecology movement (Table 1). However, Naess distinguishes four levels of deep ecology discourse constructed on and below areas of agreement within Level 2 of these platform principles (Table 2). Above, Level 1 admits the role of a developed worldview or religion. Today, deep ecology positions are developed within many traditions, including New Age Panentheism, Dharmic religions such as Buddhism and Hinduism, and Abrahamic traditions such as Christianity and Islam. Moving downward from Level 2 principles, the scheme leads through a Level 3 of policy development to Level 4, the level of practical actions, which are executed through grassroots movements as in Gandhi’s Sarvodaya model of development.

Deep ecology is a broad school of thought, but at its heart, it is always about personal transformation—a personal spiritual quest, a search for a Holy Grail of enlightenment and self-realization. Equally, deep ecology writings commonly demonstrate typically eclectic New Age influences and meta-narratives that have roots in theosophy and
DEEP ECOLOGY MOVEMENTS

1. The well-being and flourishing of human and nonhuman life on Earth have value in themselves (synonyms: intrinsic value, inherent value). These values are independent of the usefulness of the nonhuman world for human purposes.

2. The richness and diversity of life forms contribute to the realizations of these values and are also values in themselves.

3. Humans have no right to reduce this richness and diversity except to satisfy vital human needs.

4. The flourishing of human life and cultures is compatible with a substantial decrease in the human population. The flourishing of nonhuman life requires such a decrease.

5. The present human interference with the nonhuman world is excessive, and the situation is rapidly worsening.

6. Environmental policies must therefore be changed. These policies affect basic economic, technological, and ideological structures. The resulting state of affairs will be deeply different from the present.

7. The ideological change is mainly that of appreciating life quality (dwelling in situations of inherent value) rather than adhering to an increasingly higher standard of living. There will be a profound awareness of the difference between big and great.

8. Those who subscribe to the foregoing points have an obligation to directly or indirectly try to implement the necessary changes.

Table 1  The platform principles of the deep ecology movement


<table>
<thead>
<tr>
<th>Level</th>
<th>Ultimate premises</th>
<th>Platform principles movements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level I</td>
<td>Ecosophies T (T”), Buddhism, Christianity, Hinduism, Taoism, etc.</td>
<td>Deep Ecology Movement, Education for Sustainable Development Movement, Peace Movement, Planetary Citizenship Movement, etc.</td>
</tr>
<tr>
<td>Level II</td>
<td>Policies</td>
<td>A, B, C, etc.</td>
</tr>
<tr>
<td>Level III</td>
<td>Practical actions</td>
<td>W, X, Y, etc.</td>
</tr>
<tr>
<td>Level IV</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2  Deep ecology: Levels of questioning and articulation


1960s counterculture. However, despite an approach to meta-narrative that is far from post-modern, the deep ecologist’s emphasis on personal experience and emotional intelligence draws it close to the realms of cultural geography, notably situationism and its schools of landscape aesthetics. These affinities and differences are most clearly highlighted by the deep ecology doctrine of the wilderness epiphany, an individually unique threshold of self-realization that is traditionally crossed while in communion with nature in a remote place, which for Naess was the crossed stones of the Tvergastein mountain in Norway. It remains a cliche of physical geography fieldwork that it should involve long periods of living “on the edge” in a remote region, but in fact, what is going on, consciously or unconsciously, is a retreat from the world of the human and into a communion with the world of nature and identification as kinship in a Gaian sense. On a more practical Gandhian level,
Deep ecology has been proposed as a means of validation for applied physical geography.

**Fellow Travelers**

Deep ecology has many fellow travelers. Related concepts include ecopsychology, which concerns the fostering of ecological consciousness, healing the division that has grown between humans and nature and links outward into environmental psychology and social work; transpersonal ecology, which again describes a sense of self that extends beyond one’s ego and seeks an ecosophy based on identification rather than values; and interbeing, an idea that arises from engaged Buddhism and the mindfulness teachings of Thich Nhat Hanh, which flows into deep ecology through the connective practice work of Joanna Macy. Other affiliated notions include the sacred Earth ideas of many nature religions and the Hindu doctrine of reincarnation, which places a past or future human soul within every living creature. Groups such as the University of the Trees carry forward Joseph Beuys’s expanded conceptions of art into the development of connective practices designed to awaken human ecological consciousness.

Gaia is an independent scientific theory of the living Earth. However, the Gaia concept is fundamental to deep ecology. James Lovelock and Stephan Harding have shown how the science can help us link with the Gaia Superorganism. Gaian Earth systems science is, in turn, a relative of organicist general system thinking, while systems ideas are a characteristic aspect of deep ecological thought and education.

In sum, deep ecology is the realization that the ideas of “I” and “Nature” are one. The connective practices of deep ecology concern the creation of a deep personal bond with the living world. Deep ecology regards anthropocentrism as the cause of the planet’s problems and, hence, considers “ecological self-realization” to be the most important goal for both an individual and all society.

**Martin J. Haigh**

**See also** Ecofeminism; Ecological Justice; Environmental Ethics; Gaia Theory; Nature; Nature-Society Theory; Situated Knowledge

---

**Further Readings**


---

**DEFORESTATION**

One of the most notable processes by which humankind has affected the Earth’s surface is via the conversion of forested land for agriculture, logging, and urbanization, that is, deforestation. Up through the latter decades of the 20th century, deforestation had accelerated and come to be almost totally concentrated in the tropics, almost
wholly within developing countries. In the 1990s, the extent of old-growth forests was approximately one fifth of their original cover.

Estimates of land use change over the 20th century (the lion’s share of which was concentrated in its last four decades) are crude, but research suggests that though Asia and Africa had a slightly greater percentage of their forests cleared (34% to 38%) compared with Latin America (approximately 28%), a far greater absolute amount was cleared in Latin America due to the vastness of the Amazon basin. A continuation of those rates would mean the end of tropical forests within 50 years and therefore the elimination of an area prized as a repository of biodiversity.

More recent estimates of global deforestation rates reveal that while the conversion of forest to agricultural land continues at the alarming rate of roughly 13 million ha/yr. (hectares per year), countervailing forces of afforestation (the conversion of open land into forest) via planting and natural expansion have significantly reduced the net loss of forested area. From 2000 to 2005, there was an estimated net decrease of 7.3 million ha/yr., a rate substantially lower than the average 8.9 million ha decrease per year between 1990 and 2000. During the 2000 to 2005 period, while the largest net forest losses occurred in Africa and South America, net forest loss also took place in Oceania and North and Central America. During the same period, Asia reported a net gain, primarily because of reported afforestation in China, despite the fact that Indonesia had one of the world’s highest forest-clearing rates, though exact rates are contentious. Of all forested areas, primary forest constitutes 36%. Each year 6 million ha of primary forest are lost or modified, continuing trends reported in the 1990s. Although the trend in primary forest was toward a net loss, a few European countries and Japan reported gains in primary forest.

Forest-clearing “hotspots” in the humid tropical forest biome are revealed by the fact that 55% of the clearing occurs in 6% of the tropical biome’s total area and is largely confined to the current agro-industrial clearing centers of South America and insular southeast Asia. The next level of intensity of clearing, constituting 40% of the clearing within the biome, is dispersed over 44% of the land area. The remaining 5% of clearing takes place within the predominately intact forest, which makes up 35% of the biome area, and within the 15% of the biome area that had already been substantially deforested prior to 2000. In this period, Brazil and Indonesia stand out as centers of humid tropical forest loss, with Brazil constituting 47.8% of all clearing and Indonesia another 12.8%. Other Latin America hotspots include Northern Guatemala, Eastern Bolivia, and Eastern Paraguay, with the Asian countries of Malaysia and Cambodia, particularly along the Thai border, having exceptional clearing rates. Africa’s humid tropical forest loss is primarily through low-intensity selective logging, without the agro-industrial-scale clearing seen elsewhere. The loss of humid tropical forest in Africa, therefore, constitutes a much less substantial portion of the loss of this biome, at 5.4% of the worldwide loss within this biome between 2000 and 2005.

Explanatory Frameworks

Attempts to explicate deforestation outcomes have moved beyond simple Malthusian assumptions in an attempt to capture the dynamic and complex multiscaled causal mechanisms, often framing causation in a host of proximate and underlying causes to land use and land cover change. The tropical deforestation literature seeking to categorize proximate causes has highlighted three essential types: agricultural expansion, timber extraction, and infrastructure development. The underlying causes have more often identified demographic, socioeconomic, technological, political-economic, and environmental factors.

Within geography, deforestation has been examined from a variety of perspectives, from a more purely land cover standpoint seeking improved methods of measuring its extent and rates of change across different locations to human-environment research seeking to characterize the drivers of deforestation. One of the challenges of addressing deforestation is having accurate measures of the extent, intensity, and pattern of the deforestation. Thus, many of the tools that fall within the domain of geography, such as remote sensing and geographic information science, seek improved methods for characterizing the changes in this land cover in space and time.
A recent academic framework attempting to incorporate both agency and structure for explicating land change, often via the case study, is land change science (LCS), which serves as a foundational element in the study of global environmental change and sustainability science. The focus on the individual or the “agent” of change arises from the cultural ecology and risk hazards traditions, which have focused on the behavior of local agents. Structural explanations arise from the political ecology and vulnerability traditions. While in the past each perspective has often been employed at the exclusion of the others, more recent approaches have sought to incorporate both perspectives.

**Causes**

The most common proximate cause of forest clearing on the planet is the expansion of small farms along forest frontiers, followed in importance by in situ agriculture and livestock expansion, timber harvesting for fuel and construction, and the expansion of infrastructure. Road building for the last two activities often precedes expansion into the external frontier, while the in situ agriculture and pasture expansion often follows it. These proximate, immediate causes most often have at their root demographic, political, economic, and environmental processes.
While population and demographic dynamics are invoked in the great majority of deforestation case studies, they are never the only causes but instead correspond to other proximate and underlying drivers, such as socioeconomic, technological, and ecological conditions. In the early stages of development, for example, impacts on the forest are low and then accelerate during development, followed again by lower forest impacts. From a political perspective, wealthier and more democratic countries are more likely to have stable or expanding forests when compared with poor and despotic countries, which tend to experience rapid forest loss.

Much research on tropical deforestation has focused on Latin America, which is home to the greatest area of closed tropical forests in the world. In this context, small-farmer agricultural expansion along forest frontiers is the primary proximate cause of forest clearing, and examples of rapid forest conversion following colonization are abundant in the literature. Satellite imagery has illustrated particularly high rates of clearing adjacent to new roads. Road building, therefore, is a prerequisite to frontier deforestation, pointing toward corporate and state policies, which promote access to certain regions for economic or geopolitical motives. These rural frontier environments are characterized by their remoteness, abundant but often insecure resource access, and scarcity of labor. This farming technique often leads to the depletion of soil nutrients from the thin, oxidized tropical soils in 2 to 4 years’ time, leading to farm abandonment on completion of a swidden farming cycle, followed by the consolidation of lands in the hand of rural elites, who often convert the land to cattle pasture. The poor then often leave to go to the next forest frontier, beginning this deforestation cycle anew. An often overlooked point in the literature is that demographic, ecological, and political-economic drivers elsewhere spur migration to the frontier and thus serve as a necessary underlying cause of frontier agricultural expansion.

While much frontier deforestation has, at least as the proximate cause, been at the hands of small subsistence farmers, a recent trend in the expansion of large-scale mechanized agriculture into frontier-forested areas has introduced an alternate pathway to deforestation. Frontier areas of the Amazon, for example, have seen a rapid rise in the establishment of industrial-scale plantings of crops such as soybeans, though these plantings are still dwarfed in areal extent by more traditional frontier uses of small-scale agriculture and pasture. The Brazilian state of Mato Grosso, for example, saw a trend in the rise of large-scale deforestation events for the establishment of industrialized soybean farms from 2001 to 2004, often in direct relation to the price of soy in global markets. Given the current trends of soy farmers in the developed world converting to cropping maize to meet the production demands for biofuels, the forest frontiers of the developing world are increasingly vulnerable to this industrial cropping.

Unlike in Latin America, a greater proportion of deforestation in Africa derives not from colonization along the agricultural frontier but from the expansion of sedentary intensive agriculture and fuelwood harvesting. Wood fuel demands have a great impact on deforestation in Africa, perhaps more than farm expansion, though this is not always the case in countries with significant remaining tracts of forest. In the humid forests of Central Africa, logging concessions have become the most extensive form of land use in the area, making up 30% of the area, and, even though selective, alter the ecosystem and increase depredation in affected areas.

In Asia, the principle factors in deforestation have been commercial logging and the continued migration of peasants into formerly remote areas on roads established by loggers. As in all world regions, scale-dependent relationships representing social, biophysical, and geographical factors interact dynamically. Results from a case study in Thailand indicate the importance of population factors at finer scales and biophysical factors at coarser scales for explaining the variation in plant biomass levels.

**Impacts**

Even though deforestation is largely limited to within the tropics, its impacts are global, affecting global biogeochemical cycles and hydrological flows and threatening to aggravate climate change at both the local and global levels. The geographic literature has placed an emphasis on highlighting the spatial variation of the environmental impacts.
of each of these processes, such as the disproportionate warming of the tropical and subtropical climates from forest clearing. Similarly, geographic research has shown not just the total forest area lost but also that the patterning of deforestation has profound implications for the physical landscape, with forest fragmentation leading to the inhibition of forest regrowth and biodiversity loss and compromising the integrity of ecosystems.

Geographers also play a lead role in the understanding of how deforestation affects the biological integrity of tropical ecosystems. Despite the fact that this biome covers only 7% of Earth’s terrestrial surface, virtually all the recent species extinctions occurring in recent years have taken place in it, causing damage to Earth’s biological gene pool, a potential treasure trove for science, medicine, and food advancements.

A disproportionate number of global species extinctions are concentrated in those places set up to protect them, such as national parks and ecological reserves. The amount of land under legal protection has increased at an exponential rate, and by 2005, 17.1 million square kilometers of total land across more than 100,000 protected areas had been established, constituting 11.5% of the planet’s terrestrial surface. As governments expand the area of wild lands under protection and as little unclaimed forest remains outside these areas, protected areas represent an increasingly large proportion of unoccupied land available to migrant farm households.

**Conclusion**

The study of deforestation is quintessential geography. Given the enormity of the process to humans and the environment, geographers stand to make an ever-greater contribution, being strategically positioned to pioneer future research endeavors by pressing the discipline’s comparative advantage of linking human-environment interactions at diverse spatial scales. Future efforts within geography include increasing the efficiency and accuracy of monitoring via remote sensing. And given the emphasis on the case study within land change science, developing further methods of comparability between case examples will likely increase in importance.

*David L. Carr and Laurel Suter*

*See also* Biodiversity; Biome: Tropical Rain Forest; Forest Fragmentation; Global Environmental Change; Human Dimensions of Global Environmental Change; Land Use and Cover Change (LUCC)

**Further Readings**


Deindustrialization refers to the large-scale loss of manufacturing jobs, a process of capital abandonment that generally leaves devastated communities in its wake. The process reflects the mobility of capital, its constant and restless search to minimize costs and maximize profits, and its mounting ability to pit places against one another and force workers to make concessions in wage rates and benefits. The traditional view is that deindustrialization in some places and industrialization in others are mirror images of each other. Industrial growth and decline are offsetting tendencies, representing a zero-sum, or even a positive, global game. The shift of production processes from the industrial heartland to the periphery releases a skilled labor force for more sophisticated forms of production in developed countries and allows labor in the developing countries to move from relatively unproductive employment to more highly productive employment in industry. The shift may lead to some transitional unemployment, but job losses in the industrial heartland are of little significance compared with the enormous rewards attached to a global reallocation of production. Typically it witnesses steep declines in manufacturing as a proportion of the GDP and the labor force (see Figure 1).

The process of deindustrialization, often held to be synonymous with the collapse of Fordist systems of production in the late 20th century, began in earnest in the 1970s, especially in Britain and the United States and then in numerous other industrialized countries as well. In Britain, the decline of textiles, iron, and steel production, coal mining, and automobile firms devastated many once-prosperous cities of Northern England and Scotland.

Several factors contributed to the deindustrialization of the United States. Between 1945 and 1960, most United States–based companies were content to produce in the old industrial districts. But by 1960, Western European countries and Japan had become competitors. Mounting international competition and falling profit rates at home coerced American companies to decentralize not only within the United States but also abroad. Thus domestic restructuring and internationalization can be seen as two sides of the same coin. By 1980, the 500 largest U.S.-based corporations employed an international labor force almost equivalent to the size of their labor force within the United States. The rise of foreign competitors, many with lower labor costs, state-of-the-art technologies, and public subsidies, led to burgeoning trade deficits in many critical sectors, especially steel, automobiles, textiles, industrial parts, tool and die, agricultural and medical equipment, and merchant ships. U.S. corporate reinvestment, including research and development, tended to lag behind its competitors, as did productivity growth. Others note that technological change and capital intensification would have generated manufacturing job loss even without mounting international competition.

The United States experienced massive industrial devolution in the 1970s and 1980s, a period during which its share of world manufacturing output decreased significantly. This trend was accompanied by the growth of manufacturing employment and output in other countries, especially the newly industrializing countries in Eastern Asia. Inside the United States, a major corporate crisis led to declining rates of profit, inducing ever-more mobile firms to switch capital in space, going global in an effort to restore profitability: 500,000 manufacturing jobs per year were lost between 1978 and 2002. Toy production, for example, has largely moved into East Asia, particularly China. The textile industry fled much of the South for cheaper sites in Taiwan, Hong Kong, Singapore, and, now, China. The automobile industry, caught by the petrocrisis and cheap, fuel-efficient imports, gradually gave way to competitors from Japan and Europe. Steel offers another compelling example of the decline in industrial capacity. Between 1970 and 2005, the North American and Western European proportion of total global steel production declined from 67% to 42%, whereas developing countries’ production levels increased from 10% to more than 30%. Many steel-manufacturing firms have gone out of business as the global steel production capacity exceeds global demand. Because of government subsidies, steel mills in some countries, especially in Europe, have remained open in the face of dwindling quotas. The U.S. government,
however, has been less willing to pay unemployment compensation to displaced workers and has allowed the U.S. steel industry to decline. Since the 1970s, U.S. production has decreased by 33%, whereas employment in the steel industry has declined by 66%.

The dispersal of manufacturing investment to foreign lands resulted in enormous savings for American firms and enormous losses for American workers, largely due to the differences in wage rates. In terms of average wages, workers in Europe, particularly Scandinavia and Germany, fare the best, with wages often one third higher than those of American workers (as well as longer paid-vacation periods and universal health coverage). U.S. wages are comparable with those of Australia, slightly higher than those of Ireland or Japan, and well above those of the developing world.

The effect of the globalization of manufacturing change has been most severe in the Manufacturing Belt, which lost millions of relatively well-paying jobs as factories closed, with devastating economic and social impacts on their communities. Indeed the Manufacturing Belt was effectively transformed into a “Rust Belt” through a painful process of industrial restructuring, which led to numerous corporate bankruptcies, plant closures, layoffs, and rising unemployment. The region contends with problems of obsolescence and reduced productivity, especially in leading industries such as steel, tool and die production, agricultural implements, rubber, automobile and auto parts manufacturing, and shipbuilding. Many of its inner-city areas are littered with closed factories, bankrupt businesses, depressed real estate, and struggling blue-collar neighborhoods. Faced with massive job losses and eroding memberships, many unions succumbed to corporate demands for wage and benefit rollbacks. Indeed, the loss of high-paying manufacturing jobs is one of the major reasons for the stagnant incomes of American households since the 1970s. As the multiplier effects spread throughout local areas, retail sales, personal services, and real estate markets suffered accordingly. Communities increasingly desperate to prevent plant closings readily negotiated tax breaks and other incentives to keep companies from leaving town. The effects of disinvestment on workers and their communities have been devastating. Victims of plant closings sometimes lose

![Figure 1](image-url)  
**Figure 1** Decline in manufacturing as share of U.S. labor force. Deindustrialization led to significant declines in manufacturing jobs and output.

*Source: Author.*
not only their current incomes but often their accumulated assets as well. When savings run out, people lose their ability to respond to life crises and often suffer bankruptcy, depression, marital problems, homelessness, and suicide. Although job losses occurred in many occupations, some groups were more vulnerable than others. Unskilled workers are particularly likely to bear the costs of globalization, including job displacement. African American workers, many of them unskilled or semiskilled, were particularly hard hit, and by driving up unemployment, the deindustrialization of the inner city was in no small part responsible for the creation of the impoverished ghetto communities there.

Although the widespread manufacturing decline has produced a lasting effect on people and communities in the Manufacturing Belt, all is not lost in the region. There have been numerous attempts to respond to the economic crisis. Some old industrial cities such as Pittsburgh successfully built new bases for employment in services, and others, such as Cleveland, indicate the potential for doing so. Southern New England, which suffered high unemployment rates throughout much of the post–World War II period, underwent a new round of industrial expansion based on electronics and producer services, one that took advantage of its pool of highly skilled workers. Selective reindustrialization and the migration of manufacturing from the American Manufacturing Belt to the South and the West helped disguise, but not alleviate, the trauma of the Midwest and Northeast.

Barney Warf

See also Flexible Production; Fordism; Industrialization

Further Readings


A delta is the landform that results where a sub-aerial channelized flow (e.g., a river) enters a basin of standing water (e.g., the sea, a lake). In deltaic areas, flow velocities drop and sediment transport stalls, which results in bifurcation of rivers, a downstream widening of floodplains, and the buildup of deltaic plains that make the coastline protrude. Many deltas have become densely populated today. Industrial and urban uses of deltaic areas must deal with issues of flood hazard mitigation.

The word *delta* resembles the character “Δ” in the Greek alphabet. Its shape mimics that of the largest delta known in ancient Greece (that of the Nile River in Egypt as viewed from the north). Deltas come in many forms, and the term is used in various contexts. Fluvial landforms such as river channels, levees, and floodplains; lagoonal landforms, including swamps, wetlands, and lakes; and coastal features such as beach ridges, spits, salt marshes, mangrove, tidal creeks, and tidal rivers (estuary channels) can all be part of a larger area typified as a “delta,” which can further contain isolated “islands” of inherited, partly eroded, and/or buried older landforms.

Deltas grade upstream into feeder valleys, with no strict definition of their inland limits. At their downstream end, deltas terminate in open water and grade laterally into nondeltaic segments of coast—again without strictly defined limits. The word is applied to describe features at a range of spatial scales, from the small size of a flume experiment or a mouth bar at the end of a natural channel (typical in engineering textbooks) to the much larger area occupied by a full network of divergent river channels between the most-upstream bifurcation (“delta apex”) and the most-downstream subaqueous deposition as can be traced offshore (typical in geomorphological textbooks) to even larger areas based on the presence of deltaic deposits that a river has delivered over geological time over thousands (Holocene) if not hundreds of thousands (Quaternary) or millions of years (Cenozoic) (typical in sedimentary geology textbooks).

For example, the Netherlands as a whole is geologically considered a Quaternary delta of the
Rhine and Meuse within the shallow southern North Sea basin in northwest Europe. But the central part of the Netherlands is also geographically considered the Holocene Rhine-Meuse Delta. Furthermore, areas where the deltaic Rhine branches terminate in the Netherlands’ central lagoon host small deltas, which are considered elements of the larger Rhine-Meuse delta. The urbanized westernmost part of the delta, Randstad Holland, a conurbation along the mouth of the rivers Rhine and Meuse (Rotterdam), the coastal barrier (The Hague) and the central lagoon (Amsterdam), is commonly referred to as a “delta metropole.” The Randstad conurbation covers the downstream part of the Rhine delta and adjacent coastal and lagoonal landforms. In the Rhine delta, the raising of dikes has restricted deltaic sedimentation in the past 800 years to narrow embanked floodplains. Consequently, the modern floodplain has a width that is only a fraction of the preembanked situation—with impacts on style of sedimentation (allowing the floodplains to be mined extensively for bricks) and river management. In contrast, the dike-protected flood basins were starved from sediment and became subject to land subsidence owing to artificial water table lowering—a common phenomenon and management problem in urbanized deltas.

Deltas are progradational landforms that build out in size and shape up to a dynamic equilibrium state where coastal erosion balances river sediment supply—provided that circumstances such as water level (base level) and climate are semistable for a sufficiently long time. For modern coastal deltas, this has been more or less the case for the past 6,000 years, and the history of human occupation of modern deltas goes back to that time. The various types of deltas that result today are typically classified based on the physical conditions in the receiving basin (deep/shallow, tides, wave regime), character of the feeding river (sediment capacity and type of sediment, discharge regime, multichannel braiding or single-channel meandering feeder), and inherited geographical setting (nature and geometry of the basin, climate, degree of urbanization). Deltas produce fan-shaped bodies with sedimentation up to a water level, and for that property “deltaic fans” are distinct from “alluvial fans” (fan-shaped bodies above a water level as found along the foot of mountain chains) or submarine “abyssal fans” (fan-shaped bodies as found at depth on the ocean floor along the rims of continents).

“Estuaries,” like deltas, are landforms that result where rivers enter a water basin. The
difference is that estuary principally refers to the water body, that is, the river’s mouth, and an estuary experiences an inland marine influence (drowned valley, tides, salinity gradient), whereas delta principally refers to the body of accumulated sediment that protrudes into a water body (see photo). This conceptual difference ensures that the two terms are not mutually exclusive. In practice, whether an area surrounding the mouth of a river is named a delta (with or without channels termed estuaries penetrating it) or estuary (with or without substantial deltas filling it) reflects the topography surrounding the river mouth, the size of the river mouth water body versus total delta/estuary width, and whether or not the coastline protrudes into the basin in the direct vicinity of the river mouth. Estuaries of larger rivers often fill up from their upstream end by a deltaic deposit, for example, the Tagus (Tejo) River in Portugal and the River Paraná building out into the Rio de la Plata in Argentina.

Deltas often host great amounts of sedentary and migratory birds, fish, aquatic mammals, and so on, owing to a rich diversity of spatially connected ecotopes. These delta ecotopes differ in frequency, duration, and nature of flooding. Deltaic ecotopes have a relative high level of nutrient availability due to (a) a constant fresh nutrient supply by floodwaters (dissolved and as a fine sediment) and (b) waterlogged soil conditions that promote biological recycling. Marine deltas furthermore have a salinity gradient with distance to the river mouth, adding another ecological gradient and further diversity.

In the modern urbanized world, many deltaic areas have become economic hotspots, hosting a vast majority of the global population (60% to 80%, depending on one’s definition). This is the outcome of a complex history of occupation (settling, transport, urbanization) and economic development (trade, imperialism, globalization). The fringes of deltas were originally attractive settling sites because of access to water, wood, and food and opportunities for transportation. Flooding occurs frequently in natural deltas, both from rivers and from the sea, but paradoxically the flood risk is low; the sheer width of the flooded area makes extreme floods away from the coast and river cause only marginally higher water levels compared with normal floods. Consequently, living on modestly raised structures provides fair protection against floods of smaller and larger magnitudes. This changes with increased urbanization and introduction of embankments in historical (e.g., the Po Delta, Italy; the Rhine Delta, the Netherlands) to modern times.

Where delta populations and their spatial needs continue to grow after the most suitable areas in the delta have been occupied (e.g., natural levees and beaches where sedimentation occurs up to the highest flood levels and that thus flood last and least), less suitable areas such as wetlands are occupied. At that stage, embankment becomes necessary, compartmenting the floodplain and reducing the flooding frequency outside restricted dedicated flooding areas (retention basins). Embankment prevents regular flooding but increases potential damage; as a result, when flooding does occur, it has severe impacts (e.g., Hurricane Katrina). When the delta is almost completely urbanized, in addition to reusing already occupied space in new ways, there are three other strategies to further urbanize an area: (1) using the vertical dimension (building skyscrapers, tunnels), (2) reclaiming coastal water to land (building into the sea), and (3) expanding inland into the feeding valley or surrounding uplands (at growing distances to existing centers of economic activity).

Given the projected sea-level rise due to global warming in the next century, in combination with an increasing population and urbanization, many deltas in the world face severe planning issues in mitigating the flood hazard while maintaining economic vigilance and the quality of nature and living.

Kim M. Cohen

See also Coastal Erosion and Deposition; Coastal Hazards; Floodplain; Floods; Nature-Society Theory; Rivers; Urban and Regional Development

Further Readings

DEMOCRACY

Democracy means “rule by the people.” The meaning of the people and of rule is far from straightforward. Delimiting the scope of rule and of the identity of the people are intrinsically geographical processes. Since the 18th century, the normative assumption of much democratic theory has been that the people, or demos, is coterminous with the identity of the nation and that the territorially bounded nation-state is the primary agent of democratic rule and subject of democratic legitimacy. Research on democracy in geography falls into two broad areas: research in electoral geography, which investigates the mechanisms of liberal-representative democracy, and research in critical human geography, which focuses on alternative sites and spaces of democracy.

Electoral Geographies of Liberal-Representative Democracy

Liberal democracy refers to forms of institutionalized popular representation, involving periodic mass election of representatives to authoritative legislatures, under conditions of free speech and association. This model of democracy is unevenly developed in the West and is often presented as the ideal to be emulated throughout the world. Electoral geography focuses on how the mechanisms of representative politics are spatially organized in liberal democracies. This field maps the spatial distribution of votes, explains the context-specific factors affecting voting behavior, and explains how the spatial organization of electoral systems affects how votes are translated into representative majorities in liberal democracies. Research on electoral processes has also broadened out to include the geographies of campaigning, party formation, and political communication. It has also focused on the processes of democratization. The so-called diffusion of democracy as a global form of governance since the late 1980s has followed in the wake of the collapse of communism in Eastern Europe, political transitions away from authoritarianism in Latin America, Africa, and Asia, and the application of norms of democratic governance in the geopolitics of Western international financial policy, trade negotiations, and military engagements. Geographers have investigated whether the adoption of democratic forms of governance can be accounted for by specifically geographical factors. They have also critically assessed the theoretical assumptions and the practical devices through which liberal forms of electoral democracy have been circulated as the global norm.

Critical Geographies of Democracy

Democracy has only recently become an explicit object of concern in critical human geography. Recent work in this tradition searches for the signs of alternative understandings of radical democracy in the fractures and margins of liberal-representative polities. This work defines democracy as more than a set of procedures for legitimizing the decisions of centralized bureaucracies and holding elected representatives accountable. Radical democracy is understood to be a process of ongoing contestation, in which the objects and subjects of politics are constantly redefined. Radical democracy holds
to an alternative sense of “democratization,” understood not as the geographical diffusion of established norms of democracy but as the deepening of democratic impulses and their extension of new arenas of everyday life. Radical democracy also opens up to scrutiny the role of all sorts of social and cultural practices in sustaining or undermining a broader democratic culture through their contribution to the quality of the public sphere.

Geographers working on radical democracy tend to be suspicious of the normative value ascribed to the nation-state in liberal democratic theory and practice. There is a burgeoning literature that focuses on the city as the privileged scene for realizing the possibilities of radical democracy. However, this work easily falls into the trap of assuming that subnational scales of governance are somehow more democratic by virtue of being closer to people’s everyday concerns. There is also increasing attention given to emergent forms of transnational democracy. This work focuses on whether systems of globalized economic and political governance can be subordinated to democratic oversight. Settled assumptions about the objects of democratic decision making are challenged by the proliferation of nonterritorial concerns, such as global climate change, pandemic disease, and integrated transnational financial markets. These sorts of issues throw into relief the limitations of territorialized models of democratic representation.

Theorizing Democracy Geographically

To date, geographers have largely relied on normative theories of democracy drawn from other fields, political science, political theory, and political philosophy. They have applied and assessed these theories empirically, and they have often argued that they suffer from inadequately complex understandings of space, place, and scale. Meeting the challenge presented by David Slater, of developing a geographically informed conceptualization of democratic politics that can talk back to these disciplines, has proven more difficult. There are three aspects to an agenda for theorizing democracy geographically that would respond to this challenge:

1. A charitable interpretation of the imaginary geographies of democratic theory: Rather than castigate other disciplines for not holding to the spatial concepts that geographers have perfected, this agenda is sensitive to the analytical and normative problems that democratic theorists are articulating when they have recourse to what, from the geographers’ perspective, appear to be rather stylized understandings of, for example, globalization, borders, or the transnational. For example, discussion of geographical boundaries serve the function of framing the problem of how to translate broadly diffused processes of opinion formation into legitimate and effective forms of will formation; or appeals to the value of the nation-state form serve as the frame through which to reflect on the qualities of social solidarity associated with democracy. If geographers aim to contribute to debates in democratic theory, then they must do more than simply point out the limitations of these spatial “metaphors” while not addressing the analytical and normative problems at stake.

2. A diagnostic investigation of the types of influence to which particular fields of power are susceptible: A key issue in any democracy is whether the forms of “influence” that can be generated in the civil society and the public sphere can or should be articulated with the institutionalized “exercise” of power. Analysis of the “steering media” through which different fields of practice are coordinated can provide resources for an analysis of the forms of influence that different practices might be susceptible to, the forms of contention and grievance they generate, and the type of democratic politics that might be expected to emerge around them. It is here that geographers’ sensitivity to the differential spatiotemporal constitution of fields of power, added to an appreciation of the differential validity claims enacted by these formations, might contribute to broader projects of theorizing democratic futures.

3. A parasitical analysis of the ordinary deployment of normative concepts of democracy in political processes: This sort of parasitical analysis focuses on how democratic norms are invoked in new contexts in the course of ongoing political contestation; it tracks the routes through which
democratic practices circulate, and the conjunctures that account for their articulation in particular contexts. By focusing on the contestation of democratic norms in historical-geographical contexts of application, such an analysis helps disclose what values are invoked, and what harms or concerns are motivating different actors, when recourse is made to the discourses and devices of democracy. This style of analysis is context-sensitive but attuned to the processes of translation through which democratic practices travel; it is attuned to different understandings of what democracy is good for; it is sensitive to the articulation of democratic practices with non-democratic practices of bureaucracy, violence, patriarchy, and so on; and it is sensitive to the ways in which new meanings accrete to “democracy” in this process of translation and contested application.

This program for geographical research on democracy does not abandon the normative dimension of democratic theory. The mobility of democratic practices suggests that the devices through which different imperatives of democratic rule are enacted can be combined, adapted and reordered in different geographical contexts. As a consequence, the values enacted through these devices—the ways in which interests are represented, conflicts resolved, participation practiced, and accountability enforced—might be highly variable between different geographical contexts. Acknowledging that democracy collects together a series of values—including liberty and equality; participation and publicity; accountability and accommodation; contingency; contestation; and consensus, responsibility, and representation—suggests that critical attention should focus on the ways in which particular claims to instantiate democracy advance certain values over others. Judgment over the validity of any such combination will, no doubt, remain open to further contestation.

Clive Barnett

See also Citizenship; Civil Society; Electoral Geography; Justice, Geography of; Political Geography; Redistricting; Social Justice; Social Movements

Further Readings


DEMOGRAPHIC TRANSITION

Developed by several demographers in the 1950s, the demographic transition stands as an important alternative to neo-Malthusian notions of population growth. Essentially, it is a model of a society’s fertility (birth rate, BR), mortality (death rate, DR), and natural population growth rates (NGR) over time, using the simple relationship $\text{NGR} = \text{BR} - \text{DR}$. Because this approach is explicitly based on the historical experience of Western Europe and North America as they went through the Industrial Revolution, “time” in this conception is a proxy for industrialization. This approach can be demonstrated with a graph of birth, death, and natural growth rates over time that divides societies into four major stages (Figure 1).

Stage 1: Preindustrial Society

In the first stage, a traditional, rural, preindustrial society and economy, fertility rates are high and families are large and extended. In agrarian economies, children are a vital source of farm labor, helping plant, weed, and sow crops, tending to farm animals, performing chores, carrying water and messages, and helping with younger siblings.
Children also take care of their elderly parents when they are infirm. In societies with high infant mortality rates, having many children is a form of insurance that a significant number will survive until adulthood. Thus, the distribution of birth rates around the world reveals that the poorest societies have the highest rates in the world, particularly in sub-Saharan Africa and most of the Middle East. In contrast, birth rates in North America, Europe and Russia, Japan, Australia, and New Zealand are relatively low.

However, in preindustrial societies, mortality rates are also typically quite high, which means that the average life expectancy is relatively low. The primary causes of death in poor, rural contexts are the result of inadequate diets, as well as unsanitary drinking water and bacterial diseases. The world geography of death rates thus closely reflects the wealth or poverty of societies. Because both fertility and mortality rates are high, the difference between them—natural population growth—is relatively low, fluctuating around...
DEMOGRAPHIC TRANSITION

DEMOGRAPHIC TRANSITION

zero. While relatively few societies in the world live in these circumstances today, Stage I may describe certain tribes in parts of Central Africa, Brazil, or Papua New Guinea, that is, the places most remote from the world system.

Stage II: Early Industrial Society

The second stage of the demographic transition pertains to societies in the earliest phases of industrialization, such as 19th-century Britain or the United States, or selected countries in the developing world today, such as Peru. Early industrial societies retain some facets of the preindustrial world, particularly high fertility rates. Because most people still live in rural areas, children remain an important source of farm labor. The major difference is the decline in mortality rates, which leads to longer life expectancies. Mortality rates decline as societies industrialize, not primarily because of better medical care but because of improved food supplies due to the industrialization of agriculture, which plays a major role in improving immune systems, including lowering infant mortality rates. Because the death rate has dropped but the birth rate has not, the natural growth rate grows explosively, a situation evident in a wide number of countries in the developing world today.

Stage III: Late Industrial Society

Societies in the throes of rapid industrialization, in which a substantial share, if not the majority, lives in cities, exhibit a markedly different pattern of birth, death, and growth rates compared with those earlier in the transition. Death rates remain relatively low, but in this stage fertility rates also exhibit a steady decline. Birth rates typically fall and families get smaller as societies become wealthier because urbanization and industrialization change the benefit/cost ratio of children. In societies in which large numbers of women enter the paid labor force—become commodified labor outside the home, rather than unpaid workers inside it—mothers must typically drop out of the labor market, if only temporarily, to take care of their children. Economically, this process generates an opportunity cost to having children: The more children a couple has, or the longer a mother refrains from working outside the home, the greater the opportunity cost she faces, or she and her husband face as a family. As women’s incomes rise, either over time or comparatively within a society, the opportunity cost of children increases accordingly, leading to lower fertility rates. As fertility rates decline, so too does the natural growth rate. In short, relatively prosperous societies tend to have smaller families, and there is frequently a corresponding shift from extended to nuclear families in the process.

Historically, fertility levels fell first in Western Europe, followed quickly by North America, and more recently by Japan and then Eastern Europe and Russia. In these areas, reproductive levels are near, or even in some countries below, the level of generational replacement. Elsewhere, however, birth rates remain at much higher levels, although in China and southeast Asia the birth rates are dropping quickly. There has been a modest decline in South Asia, the Middle East, much of Latin America, and parts of sub-Saharan Africa.

Stage IV: Postindustrial Society

The fourth and final stage of the demographic transition, postindustrial society, portrays wealthy, highly urbanized worlds, a context indicative of Europe, Japan, and North America. Such societies typically witness low death rates, the causes of which change from infectious diseases to lifestyle-related ones, particularly those associated with smoking and obesity, as well as, to a lesser extent, car accidents, suicides, and homicides. Birth rates too, continue to fall in such contexts, as many couples elect to go childless or have only one. When birth rates drop to the levels of death rates, a society reaches zero population growth (ZPG). When birth rates drop below death rates, as they have in most of Europe, Russia, and Japan, the society experiences a negative natural population growth. Such countries are characterized by large numbers of the elderly, a high median age, and a relatively small number of children, all of which have dramatic implications for public services.

Globally, uneven economic development generates uneven patterns of natural population growth. The most rapid rates of increase are found throughout the poorer parts of the developing world, that is, sub-Saharan Africa, the Arab
and Muslim worlds, India, and Indonesia. The economically developed nations, in contrast, have low rates of growth, including in North America, Japan, Europe, Australia, and New Zealand.

**Criticisms of Demographic Transition Theory**

Although the concept of demographic transition has wide appeal because it links fertility and mortality to changing socioeconomic circumstances, it has also been criticized on several grounds. Some critics point out that it is a model derived from the experience of the West and then applied to non-Western societies as if they are bound to repeat the exact sequence of fertility and mortality stages that occurred in Europe, Japan, and North America. It is not inevitable that the developing world will follow in the footsteps of the West. Some have pointed out that the developing world is in many ways qualitatively different from the West, in no small part because of the long history of colonialism. Furthermore, demographic changes in the developing world have been much more rapid than in the West. Whereas it took decades, or even centuries, for mortality rates in Europe to decline to their modern levels, in some developing countries the mortality rate has plunged in only one or two generations. Because mortality rates do not vary geographically as much as fertility rates do, most of the spatial differences in natural growth around the world are due to differences in fertility.

*Barney Warf*

*See also* Development Theory; Fertility Rate; Malthusianism; Mortality Rate; Natural Growth Rate; Neo-Malthusianism; Population Geography

**Further Readings**


---

**Dendrochronology**

Dendrochronology, a name derived from the Greek words for “tree” and “knowing the time,” is a set of techniques by which the annual growth layers of trees, called tree annual rings, can be assigned to the specific year of their formation. The history of changes in the tree’s environment can be reconstructed using various properties of annual tree rings, for example, their width, cell size, wood density, trace element composition, and radioactive and stable isotope ratios. Andrew Douglass established the scientific basis of dendrochronology in the early years of the 20th century in North America. Douglass, the founder of the Laboratory of Tree-Ring Research at the University of Arizona in Tucson, invented the technique of cross-dating by means of skeleton plots, that is, the process of assigning year dates to annual tree rings by cross-comparison of rings from several trees growing in the same area. Cross-dating is a technique that aids the dendrochronologist in relating groups of specimens to each other by matching ring patterns and determining the exact date for each ring (Figure 1).

Cross-dating is a simple technique to undertake, although it requires an experienced person with a microscope, pencil, and graph paper for the task to be successful. The set of fundamental principles that have been established for cross-dating include the following:

1. Trees that will be used for cross-dating produce one annual ring for each year.
2. One environmental factor, such as precipitation or temperature, dominates in limiting the annual growth of the tree.
3. The intensity of the growth-limiting environmental factor varies from year to year, and the resulting annual rings reflect such variations in their growth.
4. The growth-limiting factor is effective over a large geographical area.

Tree-ring series have been used to reconstruct patterns in the variations in precipitation, temperature, soil moisture, frequency of extreme droughts, forest fires, pest outbreaks, and a diversity of other growth-limiting factors over past centuries and, occasionally, millennia. What can be reconstructed specifically depends mainly on the
factor(s) that limits tree growth. Dendrochronology has made significant contributions to the science of geography, most specifically to physical geography. Physical geography centers on the spatial analysis of all the physical elements and processes that make up the environment, such as energy, water, climate and weather, landforms, soils, animals, plants, Earth itself, and the impacts of human activities on the environment.

Some of the earliest geographical writing on dendrochronology included investigations of tree rings as a means to understand arctic and alpine ecosystems and glacial environments. However, the first dendrochronological writings by a noted geographer, Ellsworth Huntington, were reports on investigations of giant sequoia tree rings, used by the author to understand climatic variations. Climatology is a time-honored theme in geographical studies because of the interest in the spatial variability of climates and the spatial and temporal variability of atmospheric circulations and related climate phenomena such as droughts, floods, and large storms. Seminal works in paleoclimatology such as Hays and Imbrie’s investigations of the origins of Ice Ages and Hubert Lamb’s compendia of global climate variations accelerated interest in past climate variations and their relationships to cultural and economic history and geography.

As people have become increasingly interested in Earth’s climate, dendrochronology has often been adopted by geographers and other scientists as one of the methods available for the reconstruction of past climates and the effects of past climates on the historical evolution of vegetation, water resources, and other landscape features of a region. Furthermore, by knowing the nature of past climates, probabilities of future climates are obtained by linking the knowledge of past climates obtained through the analysis of the tree-ring series to stochastic (statistical) models that generate the future likelihoods of selected precipitation regimes or temperature patterns.

The reconstruction of past climates through cross-dating analysis extends the climatic record

Figure 1  Concept of cross-dating: extending chronologies by cross-dating living and dead wood samples

Source: Laboratory of Tree-Ring Research, University of Arizona.
for a regime much further into the past compared with on-site instrumental measurements, which are often limited to less than 100 years. Consequently, a more reliable depiction of the nature of climates is available for projection into the future. The impacts of future long-term climatic changes on landscape features and the people living on these landscapes can be assessed and, when necessary and feasible, appropriate actions taken in the planning process to mitigate undesirable consequences and attain desirable outcomes.

With the availability of high-speed computing systems and improved spatial analysis through geographic information systems and other innovations, dendrochronology has often become an increasingly important tool in investigations of landscape geography. One measure of the contributions of dendrochronology to geography is supplied by the International Tree-Ring Data Bank’s online database of dendrochronologists (http://web.utk.edu/~grissino). This Web site lists dendrochronologists at 64 academic and research institutions in 20 countries worldwide with “geography” in their name. Dendrochronology sessions are also common at the annual meetings of the Association of American Geographers.

Ramzi Touchan and Peter F. Ffolliott

Further Readings


**DEPENDENCY THEORY**

Dependency theory falls within a group of neo-Marxist theories that attempt to explain development, underdevelopment, and inequalities in the global system. Dependency theory is a response to the earlier modernization theory and its failure to generate economic growth in Latin America in the 1950s and 1960s. Whereas modernization theory was the brainchild of thinkers in core countries, that is, the economically developed world, dependency theory was the product of intellectuals in the world periphery or less developed countries disillusioned by the idea of modernization. Instead of arguing that poor countries are simply further behind richer countries in terms of development, dependency theorists argued that poorer countries are exploited by richer, developed countries, or, more pointedly, that the development of European and other Western countries depended on the underdevelopment of non-Western countries. In this framework, poorer regions such as Latin America are peripheral to the capitalist industrial core, or the rich, developed countries of Europe and North America. Dependency theory was important in opening the door within geography and in other social sciences to discussions around the social construction of poverty.

The roots of dependency theory lie in the work of the United Nations Economic Commission for Latin America (ECLA) in the 1960s, under the direction of Argentinean Raúl Prebisch. Prebisch, in his “ECLA Manifesto,” criticized the international division of labor, in which Latin America provided food crops and raw materials for the industrial core and received finished goods in return. In his opinion, the continuation of this relationship would inhibit Latin America’s process of capital accumulation because of the unbalanced terms of trade. Prebisch proposed instead that Latin America industrialize, which would require protectionism and a heavy role on the part of governments. The ECLA’s proposal was not received well by Latin American governments and indeed was overly optimistic in assuming that industrialization would solve all the problems of development in the region.

The ideas planted in the ECLA then emerged in more radical forms in the late 1960s and 1970s from two main sources. In the United States, socialist writers such as Paul Baran of the leftist journal *Monthly Review* argued that developing economies were kept in stagnant positions by monopolistic corporations that controlled competition and profits and that the only way to achieve true development was to exit the global
capitalist system of monopolies and build society and economies on a socialist base. In Latin America, a group of radically critical intellectuals, including Fernando Cardoso, Enzo Falleto, and Celso Furtado, known as dependistas, developed a more radically critical set of theories of dependence that held that the rich countries’ development was the result of the destruction of poor countries’ social and economic systems through forced integration into the global capitalist economy and thus a changing of their previously viable institutions. This was a parting from the original idea of underdevelopment, which was conceived by core modernist thinkers as the state of being in the periphery before core intervention, that is, the problem to be solved by development. In the dependistas’ thinking, underdevelopment is not an original condition but an active process.

André Gunder Frank was instrumental in bringing the ideas of dependency theory to young North American intellectuals, who were in the midst of an era of protest and social rebellion at the time, as well as presenting dependency theory as a critique of modernization, which offers an internal explanation for underdevelopment, that is, that there is something inherently wrong with the culture, for example, of underdeveloped countries that keeps them from developing. Frank coined the term the development of underdevelopment to describe the process of underdevelopment. He argued that Latin American countries had been involved in a chain of dependence since colonial times, a phenomenon that core countries had never experienced; thus underdevelopment cannot be explained by modernization theory. Frank and other dependistas instead argued that colonialism was a forced changing of institutions, society, culture, and economy.

According to dependency theory, exploitation extended from the global to the individual scale. Thus, landless workers and peasant tenants were exploited by landowners who bought their goods cheaply, who then sold the goods to merchants for a profit, the chain continuing until the goods were exported to the core. Thus, the development of dependence and the development of underdevelopment are interdependent. The point is that the continuous chain of exploitation from core to periphery created a situation of developed and underdeveloped, one necessitated by the other, and was seen by dependistas as key to understanding the social construction of poverty. Dependency also restricted the social and political development of places marginalized by their relationship to the core, a situation that was further aggravated by the forces of the unfettered market, which tend to exacerbate socioeconomic inequalities.

Theotonio Dos Santos further developed dependency theory by distinguishing between three forms of historical dependence. Colonial dependence was the original form of dependence, in which the colonial or core power dominated and benefited from the extraction of raw materials and human resources from the colonized country. This form of dependence transitioned in the latter part of the 19th century to financial-industrial dependence, which was based on a system of specialized export of food products, as well as other sectors that depended on the export sector. The third historical form of dependence is technological-industrial dependence, which emerged after World War II. Dos Santos argued that the development of the industrial sector in the periphery was grossly restricted by its dependence on the export sector and an unequal balance of payments due to industry’s dependence on imported necessary synthetic materials.

Solutions to the problem of dependency are restricted by the ideas of dependency themselves. Harold Brookfield has argued that the core itself depends on the periphery for labor, markets, and natural resources for its own continued state of development and that therefore it becomes increasingly difficult to distinguish who depends on whom. Within the school of dependency theory, opinions differed as to the solutions to the problem of dependence. Some argued that reform of the global capitalist system with state intervention would be sufficient. Frank and other dependistas took a more determinist position and argued that development of exploited regions such as Latin America was more likely to take place during periods of less integration in the global capitalist economy, such as during World War II. This tendency argued that only the overthrow of the capitalist system would truly solve the problem of capitalist exploitation of underdeveloped regions.

Dependency theory was especially important because it came from the experiences of the non-Western world and broke the dominance of
Western voices in development theory; however, its relevancy as well as its applicability to policy has been heavily challenged. First, its assertion that peripheral countries cannot develop in the context of the global capitalist system was weakened by the spectacular growth of East Asian countries. Another critique of the dependency school is that it does not sufficiently include social or political contexts in its analysis, instead laying all the weight of causality on economic factors. A lack of attention to the social or cultural contexts of dependence signifies that dependence is only a material relationship when it is in fact social. This marks dependency theory’s limitations as an applicable theory of underdevelopment and poverty as well as a source of policy solutions to poverty.

*Heather R. Putnam*

**Further Readings**


**Derechos**

Straight-line winds associated with convective weather activity cause property and crop damage and loss of life each year, comparable with tornado losses. A derecho is such a strong linear wind that is not tornadic, with wind speeds in excess of 26 m/s (meters per second) (58 mph [miles per hour]). The term *plow wind* applies in the Canadian vernacular. The name, coined by University of Iowa physicist G. Hinrichs in 1888, derives from a Spanish word meaning “direct” or “straight ahead.”

Derecho winds generated convectively from a downburst cluster forms a *progressive derecho* (major axis of effects less than 400 km [kilometers]; 250 miles), whereas, on a larger scale, a group of downburst clusters produces a *serial derecho* (damage axis more than 400 km). Derechos tend to blast in linear paths fanning out along singular or multiple curved-wind fronts, known as *bow echos*, over a wide swath of land. Their duration and spatial dimensions provide data for identification.

The danger stems from the potentially large areas affected by the highly variable derecho winds. And derechos are associated with unstable bow echo systems that migrate rapidly, sometimes exceeding 50 mph or more in speed, making prediction difficult and forecasts uncertain. People can be overtaken by such a system with little warning.

Regarding their climatology, derechos can occur in any month but peak in May through July and form principally east of the 100th meridian in North America. Derechos pose distinct hazards to summer outdoor activities by swamping boats, throwing flying objects, and breaking trees and limbs. Typically, there are warm-season and cool-season patterns. May to August finds them in the region stretching from Southern Wisconsin across Illinois, Indiana, and Western Ohio in the upper Midwest and in a region of Oklahoma, adjoining Kansas and Missouri. This accounts for nearly 70% of occurrences.

From September through April, areas of activity migrate southward to Eastern Texas through Alabama, extending along the lower Mississippi Valley to the Ohio River and Southern Indiana. The causative association for this cool season is with migratory low-pressure systems. Derechos have been reported elsewhere in the world; however, data are lacking for specific assessment.

As an example, the speed of a derecho in 1998 in Eastern Wisconsin exceeded 57 m/s (128 mph): Overall, some 139 counties in nine states were affected in related storms. In August 2007, a
Desertification is a contentious term, most commonly referring to human- and climate-induced changes to dryland regions of the world resulting in a loss of their biological and economic potential. Much debate surrounds the degree to which this change is human or climate induced and whether or not the process is irreversible. Despite this contention, desertification is a significant environmental problem that has been the subject of two United Nations (UN) conventions (treaties) and scholarly attention of both physical and human geographers.

Desertification occurs when human action, frequently accompanied by less than favorable climatic trends, conspires to reduce the productive potential of land in arid, semi-arid, and dry subhumid areas of the world. Desertification occurs neither in true deserts (as it is sometimes mistakenly believed) nor in more humid environments. Desertification encompasses a wide array of anthropomorphic forms of soil degradation in dryland areas, which are often aided and abetted by lower-than-average rainfall in certain climatic cycles. These forms of degradation include soil acidification, compacting, crust ing, erosion, nutrient depletion, and salinization, as well as the impoverization of surface biomass and diversity. Such degradation is often brought on by deleterious modern and traditional crop farming practices (such as the indiscriminate use of mineral fertilizers, overly aggressive tillage, poor ground cover in the off season, and insufficient organic amendments); poor pasture management for livestock; and unsustainable forestry practices, mining, and maladapted irrigation practices (such as excessive diversion or flood irrigation without proper drainage) leading to destruction of natural floodplains or salinization.
human action, or a combination of the two. Furthermore, many scientists caution that geographical and temporal scale must be taken into account. Degradation thus may occur in some locations, but in other locations land that is considered degraded actually is recovering. These differing assumptions about the extent and severity of desertification worldwide inherently lead to very different ideas about how to combat it.

Causes of Desertification: Human Versus Biophysical

The causes of desertification are complex and interrelated, making it difficult to cite one single factor as its ultimate cause. Researchers generally fit into one of two groups: those who see human impacts as the drivers of land degradation follow the equilibrium approach, while those who see random biophysical processes as the main factors use the nonequilibrium approach. Human activities said to degrade marginal land include overgrazing, deforestation, salinization of soils due to inappropriate irrigation, and farming techniques such as field burning. Farmers with small landholdings stereotypically are seen as investing too little in soil fertility, while overgrazing is viewed to be the result of a “tragedy of the commons” effect. Industrial agriculture also puts stress on the land, particularly through the use of tractors and reliance on chemical fertilizers as opposed to biological soil fertility.

Biophysical factors influencing land degradation include rainfall, vegetation, wind and water erosion, albedo (the proportion of sunlight that is reflected back into the atmosphere rather than absorbed by the soil), slope, evapotranspiration,
and soil moisture. These biophysical factors influence each other through complex feedback loops. For example, lightly vegetated land is more susceptible to degradation through wind and rain erosion, which decreases soil moisture and in turn decreases vegetation. The fact that desertification occurs across cultures and on every continent on Earth, with the exception of Antarctica, lends credence to the importance of biophysical factors as underlying causes of desertification. As global climate change progresses, these factors will potentially gain even more importance as temperatures rise and rainfall patterns change.

Desertification Around the World

More than 40% of Earth’s surface is covered by drylands susceptible to desertification. Combined, these regions have a population of more than 250 million people. Each region faces unique challenges stemming from the complex interactions between people and their environments. Farmers practicing slash-and-burn agriculture in mountainous countries such as Honduras and Nicaragua, for example, face different challenges from those faced by peasants in remote oasis regions of China experiencing sand dune encroachment. On the other hand, many dryland dwellers worldwide are similar in that they tend to live in relatively sparsely populated, remote parts of the country, often distant from government. Affected populations differ greatly in their vulnerability to desertification, depending on the country’s level of development and dependence on subsistence agriculture.

While susceptibility to desertification is widespread throughout the world, Africa receives the most attention from both the scientific and the aid communities as its peoples are thought to be the most vulnerable and its lands the most degraded. The North American West may be at risk of widespread land degradation, but its inhabitants are not particularly vulnerable because the region’s economy is not based primarily on subsistence agriculture. Poor populations dependent on the land for farming and grazing, however, risk decreasing incomes and famine if their land becomes degraded. According to the conventional view, these populations are driven to farm or graze more intensively as soil fertility decreases to maintain yields, further degrading the land and leading to a vicious cycle where poverty and environmental degradation reinforce each other. If the land becomes so degraded as to no longer support vegetation, these populations have no recourse but to leave in search of less degraded land. Often, displaced populations end up on marginal lands on the outer fringes of urbanizing areas, where their subsistence activities once again degrade the land. Desertification thus increases pressure on the land, leaving poor populations increasingly susceptible to drought-induced famine and conflicts over land. This limited and conventional view of the impacts and causes of desertification has been widely critiqued by human geographers. Moreover, the conventional wisdom may shift as global climate change threatens prosperous industries in the developed world, such as the Australian wine industry, or leads to increased tensions in areas such as the Middle East as water sources dry up. Furthermore, changing climates may affect yields of staple grains produced in traditional bread baskets such as the North American Great Plains, while increasing demand for meat and ethanol worldwide decreases the amount of arable land used for grain production, driving up food prices and increasing global hunger.

Evolving Understandings of Desertification

While scientists and commentators have been writing about desertification for centuries, the concept first grabbed significant global attention during the Dust Bowl phenomenon of the 1930s and 1940s in the United States. Here, human mismanagement coupled with a prolonged drought led to widespread aeolian erosion, and devastation of many farming communities, on the Great Plains. While the proximate causes of this crisis were, in part, related to maladapted farming practices, subsequent scholarship has implicated agrarian capitalism in the United States.

Desertification again attracted worldwide attention in the 1970s and 1980s following major droughts in the Sahelian grasslands south of the Sahara Desert. During this period of below-normal
rainfall there was considerable concern that the Sahara Desert was expanding southward as the Sahel was desertified. While it was acknowledged that below-average rainfall was contributing to desertification of grasslands in this region, much of the blame was attributed to human causes such as overgrazing, population growth, and destructive crop farming practices. Subsequent scholarship revisited conventional understandings of desertification in the Sahel, leading to a series of reinterpretations of this process. First, there was a general acknowledgment that the role of local natural resources management as a contributor to Sahelian desertification during the 1970s and 1980s had probably been overestimated. Second, along with a greater recognition of fluctuating climatic variables (mainly rainfall) as the major drivers behind desertification, there was a greater appreciation of how resilient some forms of dryland vegetation could be when the rains eventually did return. Third, the recognition that the desertification process had been misunderstood in the Sahel, especially the role of local people, led several human geographers, and those in allied fields such as anthropology and environmental history, to examine the term’s past, from the colonial era to the present. This line of inquiry dovetailed with a broader postmodern movement that took great interest in the social construction of nature.

With regard to the specific history of the term desertification in the Sahel, both British and French scientists were concerned about desertification during the colonial era (and were influenced by the literature on the U.S. Dust Bowl). While some colonial era scientists attempted to apply scientific methods rigorously to understanding African environments (by, e.g., carrying out a range of applied and basic research projects to developing a savanna ecology model for the region), this science suffered from several problems. In many cases, European understandings of desertification in Africa were based on ecological models derived from temperate climates; an underlying belief that Africans employed backward or primitive natural resource management practices; the use of findings from very localized and short-term studies to explain broadscale environmental changes; and an unwillingness to incorporate local understanding of environmental processes into research and policy agendas. Explaining why these misunderstandings of desertification persisted over time has also been the subject of much scholarship. In the case of African governmental agencies and nongovernmental organizations, these entities had an incentive to perpetuate crisis understandings of desertification as they often resulted in more funding. Certain scientists and consultants also had a motivation to reinforce such beliefs if it meant future contractual assignments.

Combating Desertification

Following the Sahelian drought and famine in the late 1960s and early 1970s, the world awareness of desertification reached a peak. In 1977, the UN convened in Nairobi for a conference on desertification. The conference cited human factors such as overgrazing and destructive farming techniques, based on the assumption that underdevelopment and environmental degradation reinforce each other, the conference developed a plan of action to be implemented by national governments in affected countries. Recommended actions to combat desertification included the development of land use plans, awareness campaigns, and revegetation initiatives. Local people were called on to use their land and water resources more efficiently by improving rangeland and farmland management. Recommendations also included measures to improve rural development with the aim of increasing education and access to health care and decreasing rural-to-urban migration. The belief was that if desertification was not reversed, up to one third of agricultural land would be lost and food production would not be able to meet the demands of a rapidly growing world population.

By the time of the United Nations Conference on Environment and Development in Rio de Janeiro in 1992, the scientific community believed that desertification and degradation had only intensified over the past 15 years (although the Conference on Desertification’s dire predictions had not become reality). The UN thus initiated the Convention to Combat Desertification in 1994 with the aim of reinvigorating antidesertification efforts and paving the way for sustainable
development. The convention currently meets biennially and is implemented through national action plans administered by each signatory country. As of 2008, this convention had been ratified by 193 countries. Actions taken under these plans, such as those following the 1977 conference, include measures to combat soil erosion, reforestation campaigns, and efforts to reduce the incidence of bush fires. These measures are to be taken in conjunction with efforts to increase human development. Tools proposed by geographers include practicing local environmental knowledge to meet the unique demands of each community facing land degradation.

William G. Moseley and Erika S. Jermé

See also Biome: Desert; Environmental Discourse; Nomadic Herding; Population and Land Degradation; Soil Degradation; Soil Erosion; United Nations Environment Programme (UNEP); Vulnerability, Risks, and Hazards; Xeriscaping

Desert varnish is but one of more than a dozen rock coatings (Table 1) that drastically alter the appearance of rock surfaces. The better term is rock varnish because this coating occurs in virtually all environments, including alpine, antarctic, arctic, desert, periglacial, stream, temperate, and tropical settings. This paper-thin accretion is characterized by extremely high concentrations, typically more than 10%, of manganese oxides that give it a characteristic black to dark brown appearance. Clay minerals, however, make up the bulk of rock varnish along with iron oxides.

There are four general explanations for how rock varnish accretes on top of rock surfaces, but all these models reject the old idea that the constituents of varnish derive from the underlying rock. The model that has not yet been falsified is the polygenetic model of rock varnish formation (see first series of photos). This explanation combines bacterial enhancement of manganese and iron with abiotic fixation of the manganese by clay minerals. The process starts with bacteria fixing manganese on cell walls. Wetting events dissolve manganese. The desert dust supplies interstratified clay minerals, and the dissolved manganese reprecipitates as nanometer-sized fragments of manganese oxides. These tiny minerals fit into the weathered edges of clays, tightly cementing clays to the rock surface. The effect is a highly layered texture at the micrometer and nanometer scales imposed both by the clay minerals and the cementing manganese-oxides.

The most exciting development in varnish research in the past several decades is the development of varnish microlaminations (VMLs). This tool is based on 12 years of detailed analyses of more than 10,000 microsedimentary basins by Tanzhuo Liu (www.vmldatinglab.com). Working in the deserts of Western North America, Liu found that black varnish layers correspond with wet events and developed calibrations for the late Quaternary and a separate calibration for the Holocene. The second photo illustrates just one rock varnish VML sequence from Death Valley, California, formed over the past 20,000 years.

Further Readings


<table>
<thead>
<tr>
<th>Coating</th>
<th>Description</th>
<th>Related Terms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carbonate skin</td>
<td>Composed primarily of carbonate, usually calcium carbonate (CaCO₃) but sometimes magnesium carbonate (MgCO₃)</td>
<td>Calcrete, travertine</td>
</tr>
<tr>
<td>Case hardening</td>
<td>Addition of cementing agent to rock matrix material; the agent may be manganese, sulfate, carbonate, silica, iron, oxalate, organisms, or anthropogenic</td>
<td>Sometimes called a particular type of rock coating</td>
</tr>
<tr>
<td>Dust film</td>
<td>Light powder of clay- and silt-sized particles attached to rough surfaces and in rock fractures</td>
<td>Clay skins, clay films, soiling</td>
</tr>
<tr>
<td>Heavy-metal skins</td>
<td>Coatings of iron, manganese, copper, zinc, nickel, mercury, lead, and other heavy metals on rocks in natural and human-altered settings</td>
<td>Sometimes described by chemical composition</td>
</tr>
<tr>
<td>Iron film</td>
<td>Composed primarily of iron oxides or oxyhydroxides; unlike orange rock varnish because it does not have clay as a major constituent</td>
<td>Ferric oxide red staining, iron staining</td>
</tr>
<tr>
<td>Lithobiontic coatings</td>
<td>Organisms forming rock coatings, for example, lichens, moss, fungi, cyanobacteria, algae</td>
<td>Organic mat, biofilms, biotic crust</td>
</tr>
<tr>
<td>Nitrate crust</td>
<td>Potassium and calcium nitrate coatings on rocks, often in caves and rock shelters in limestone areas</td>
<td>Salt peter, niter, icing</td>
</tr>
<tr>
<td>Oxalate crust</td>
<td>Mostly calcium oxalate and silica with variable concentrations of magnesium, aluminum, potassium, phosphorus, sulfur, barium, and manganese; often found forming near or with lichens</td>
<td>Oxalate patina, lichen-produced crusts, patina, scialbatura</td>
</tr>
<tr>
<td>Phosphate skin</td>
<td>Various phosphate minerals (e.g., iron phosphates or apatite) sometimes mixed with clays and sometimes manganese</td>
<td>Organophosphate film, epilithic biofilm</td>
</tr>
<tr>
<td>Pigment</td>
<td>Human-manufactured material placed on rock surfaces by people</td>
<td>Pictograph, paint</td>
</tr>
<tr>
<td>Rock varnish</td>
<td>Clay minerals, manganese (Mn) and iron (Fe) oxides, and minor and trace elements; color ranges from orange to black, produced by variable concentrations of different manganese and iron oxides</td>
<td>Desert varnish, patina, Wüstenlack</td>
</tr>
<tr>
<td>Salt crust</td>
<td>Chloride precipitates formed on rock surfaces</td>
<td>Halite crust, efflorescence</td>
</tr>
<tr>
<td>Silica glaze</td>
<td>Usually clear white to orange shiny luster but can be darker in appearance; composed primarily of amorphous silica and aluminum but often with iron</td>
<td>Desert glaze, turtle-skin patina, siliceous crusts, silica-alumina coating, silica skins</td>
</tr>
<tr>
<td>Sulfate crust</td>
<td>Sulfates (e.g., barite, gypsum) on rocks; not gypsum crusts that are sedimentary deposits</td>
<td>Sulfate skin</td>
</tr>
</tbody>
</table>

**Table 1**  Major categories of rock coatings

Source: Author.

Liu and other researchers in different settings are in the process of developing calibrations for other regions of the world. Although there are other dating methods that have been proposed for varnish, VML is the only method that has gone beyond the experimental stage.

Subject to successful blind testing, the great power of this method is that it is now possible to determine the minimum age of exposure of the underlying rock surface. Liu and other researchers have applied VML to dating features such as alluvial fans, meteorite impact...
Varnish formation starts with bacteria that concentrate manganese (Mn) and iron (Fe), as revealed by electron microscope observations. Then, using a high resolution transmission electron microscope, it is possible to see how the manganese moves from the bacteria and into weaknesses in the clay minerals that fall onto desert rock surfaces. The manganese and iron then fit into the clay structure in a way that cements everything together onto rock surfaces. Thus, rock varnish is analogous to a brick wall, where the bricklike clays are cemented by the mortarlike oxides.

Source: Author.
craters, faulting events, landslides, stone artifacts, and ancient petroglyphs.

Ronald I. Dorn

See also Biome: Desert

Further Readings


Deterritorialization and reterritorialization processes are spatial manifestations of contemporary changes under way in the relationship between social life and its territorial moorings. The two terms were originally employed in the 1970s in the work of French theorists Gilles Deleuze and Felix Guattari, who are often associated with poststructuralism and postmodernism. Using insights from philosophy and psychoanalysis, Deleuze and Guattari developed a sophisticated understanding of capitalism, power, and identity,
which are locked in a fluid process of territorialization, deterritorialization, and reterritorialization of social structures and processes. These terms have subsequently been adopted by the social sciences and humanities, that is, geography, anthropology, international relations, linguistics, and others.

Deterritorialization has often been associated with globalization. The bonds that tied economics, politics, and culture to fixed spatial configurations such as national territories are loosened under globalization pressures. Globalization flows, suggesting mobility, are perceived as replacing the space of places, suggesting territorial fixity. In a broader interpretation, deterritorialization processes indicate a decrease in the significance of territory for social life, which is thus an unraveling of territoriality as it has been constructed during the modern era. It appears that a primarily networked organization of spatial power is replacing a primarily territorial organization of spatial power. Specifically, deterritorialization processes take aim at the global system of sovereign nation-states and their boundaries that has dominated the territorial organization of power during the modern era. Nation-states are devolving powers on two main geographical scales: upward to the supranational bodies such as the European Union, NAFTA, and the International Monetary Fund; and downward to subnational institutions such as regional governments, local councils, and development agencies. These developments are unseating the traditional role of the nation-state as the territorial container of social relations and opening up novel possibilities for the spatial organization of social relations.

From a political economy perspective, deterritorialization processes are understood in terms of the spatial dimensions of successive rounds of capital accumulation. If the previous strategies of capital accumulation largely took place at the scale of the national markets, the current ones favor global markets. Financial markets, in particular, are among the most dynamic global markets. Financial flows circle the globe at unprecedented speeds via digital telecommunications networks, unsettling national control over the economy and making national boundaries appear archaic. Manufacturing is outsourcing jobs from the developed to the developing economies as part of a strategy to compete in global scale markets. At the same time, numerous firms have adopted truly global production models that involve an integrated network of places.

Geopolitically, with the end of the Cold War the world political map has registered dramatic developments that have dismantled the geography of power that once appeared unshakable. Sovereign states have fragmented, while new ones have appeared on the map. The lines between the domestic and the foreign spheres of politics have blurred as national governments face increasing difficulties in managing crisis situations that often acquire transnational dimensions.

Cultural and social issues are increasingly playing out in the global arena rather than simply within nation-state borders. Globalization-induced migration flows and information technologies have created transnational networks of diasporic communities and have reinvigorated local and regional identities that are now enacted globally. Numerous governmental social policies and executive functions have been privatized, indicating a spatial shift from hierarchical government to networked governance.

Deterritorialization processes have raised critical questions regarding the spatiality of social relations under globalization. Some authors have interpreted deterritorialization as the “end of geography” and the emergence of a border-less world, in the sense that territory and territoriality will be of little consequence to social life in the future. In this view, deterritorialization appears as a transcendental phenomenon leading an unstoppable march to “a-territorial” social relations. Such a unidirectional understanding of globalization has poor explanatory power for the complex territorial changes contemporary society faces.

Evidence suggests that there is a reterritorialization of economic and political activity that transcends the spatial framework of the nation-state system. Reterritorialization processes are generally understood as the restructuring of territorial forms of organization of social relations, such as the nation-state. This implies the uncoupling of the exclusive links between state sovereignty and territory and the emergence of new territorial configurations beyond the scope and the scale of the
nation-state. Globalization has not led to ceaseless
deterritorialization but has created conditions for
the spatial reconfiguration and rebordering of
social relations. Geography has not ended, and
the world has not become border-less. Rather,
geography is being reorganized, and borders are
acquiring new significance.

The global reach that the sovereign state sys-
tem has achieved means that globalization flows
have to necessarily engage with territorial actors
such as nation-states. Moreover, networked glo-
balization flows do touch down in various geo-
graphical places, such as global cities and resource
rich regions that serve as network anchors, result-
ing in an archipelago-like geography. National
borders may do little to control financial flows,
but nation-states do play a role in regulating
global financial markets. Digital information
technologies may have the power to undermine a
government’s monopoly over information, but
the Chinese government has managed to regulate
domestic Internet content. Since the end of the
Cold War, the world political map has seen the
emergence of new geopolitical actors such as
the European Union, China, and other smaller
regional powers, indicating the persistence of ter-
ritorial power politics. Transfer of state powers
to sub- and supranational levels does not neces-
sarily mean the end of state sovereignty. Instead,
it can signal the emergence of a new architecture
of political territoriality with multiscalar and
overlapping sovereignties. Some identities may be
enacted at a global scale, but the emergence
of new ethnic and territorial identities points to
the salience of the nation-state as a territorial
framework for social life.

The emerging picture of the current moment of
globalization suggests that territoriality as a prin-
ciple of organization of social relations is far from
disappearing. The territorial nation-state, in its
quest to remain relevant in the 21st century, plays
an active role in the twin processes of deterritori-
alization and reterritorialization. States lose some
powers but also gain new ones. They are not fad-
ing away, but they reinvent themselves and share
some of their power with nonstate structures of
territorial power such as global cities, suprana-
tional organizations, and subnational regions.

Deterritorialization and reterritorialization
are best understood as processes that unfold
simultaneously. Geographically, some spaces
can experience deterritorialization, while others
may experience reterritorialization. Social rela-
tions do not completely lose their territorial
grounding and their boundaries before they
reterritorialize. Rather, social relations acquire
other territorial configurations and boundaries
even as they lose the previous ones. This means
that the new territoriality of social relations,
while being qualitatively different, will include
vestiges of the old one.

Gabriel Popescu

See also Borders and Boundaries; Globalization; Nation;
Sovereignty; State; Territory; Transnationalism

Further Readings

reterritorialisation: The re-scaling of urban
governance in the European Union. Urban

Ó Tuathail, G. (1999). Borderless worlds?
Problematising discourses of deterritorialization.
Geopolitics, 4, 139–154.

(dis)integration: Deterritorialization and
reterritorialization in the new wor(l)d order.
Annals of the Association of American
Geographers, 84, 381–398.

The developing world comprises a vast set of het-
erogeneous societies that includes the bulk of the
world’s population. The distinction between
“developed” and “underdeveloped” is usually
made in economic terms. The gross domestic
product (GDP) per capita and level of industrial-
ization are the key indicators used to define a
country’s level of development. Many factors
influence this variable, including a country’s colo-
nial legacy, government policies toward trade and
investment, resource endowments, and levels of
human capital, including literacy. One dimension
often overlooked is the role of geography. Factors regarding the physical, political, and human geography of a region are normally not considered among economists when looking at the potential development of certain regions of the world. This entry seeks to connect the physical geography of a given country and its challenges for economic development.

It is unrealistic and naïve to think that all countries in the world will develop in the same linear fashion as those in North America, Western Europe, and the Pacific Rim. The geography of many countries is too diverse and in many cases a limiting factor for them to take the same trajectory of development. Financial power centers fail to consider some of those limiting factors. However, countries are not forever relegated to living in poverty. Rather, they need to take different roads toward “developing” that take into consideration their unique geographical conditions.

Consideration will be given to the “3 Ds” that challenge developing countries: disease, drought, and distance to world markets. These are key geographical factors to consider when determining the developmental track of a “developing” country (see Figure 1).

If one were to look at a world map and label all the countries that could be considered “developed” as opposed to “developing,” the vast majority of the developed countries lie north of the Tropic of Cancer (23½° N). This area includes North America, Western Europe, Eastern Europe, South Asia, and East Asia. The developing world also includes the continents or regions of South America, Africa, and southeast Asia. Although there are obvious exceptions to this generalization, the trend is a significant one.

There are two significant reasons for the relevance of the Tropic of Cancer. First, the tropics lie between this latitude and the Tropic of
Capricorn, at 23½° S. In this area, there are unique geographic factors generally not found outside the tropical zone that challenge the developmental trajectory that individual countries are able to take. Second, there are very few landmasses south of the Tropic of Capricorn in the temperate zone of the Southern Hemisphere. Much more land lies in the temperate Northern Hemisphere: There is simply more temperate land in the north.

A number of features of the physical geography of a given country present obstacles to economic development. Natural hazards, including flooding, droughts, earthquakes, volcanoes, storms, hurricanes, diseases, illnesses, and pests can all undermine economic development. Large natural disasters can set countries back greatly in their economic development, as in the periodic flooding of Bangladesh. Volcanoes and floods can often have both positive and negative effects as they bring in fertile sediment. Areas around volcanoes and flooding deltas are often heavily populated, as in Egypt, Bangladesh, and Indonesia. Diseases such as malaria that thrive in tropical climates and AIDS, which is endemic in Africa, prevent people from working and create an economic burden on society. Pest such as locusts reduce agricultural output and make it more difficult for a country to earn sufficient money to escape from subsistence agriculture. Reliable sources of water are necessary for productive agriculture and to a lesser extent for industry.

Most people in the economically “developed world” cite population growth in poor countries as the reason for poverty and for other global issues such as illegal immigration, climate change, and economic stresses in their own countries. There is a direct relationship between an exponential population growth and the incidence of poverty. Actually growth rates and poverty have a symbiotic relationship: A rising population growth leads to more poverty, and more poverty leads to population growth. The poor compensate for their lack of resources by having more children to help with farming needs and bring in another income and to compensate for high infant mortality rates. However, the rising population means more mouths to feed in the family and fewer resources to pass on to the next generation. It is no coincidence that in the developing world it is the countries that have the highest annual population growth rate that are also those with the highest percentage of people living in extreme poverty. However, one must also consider the certain geographical factors that come into play in contributing to this cycle. Geographers see issues of the developing world not merely as economic or demographic problems that need solutions from the “developed” world. They see the issues in the totality of their context, one which includes some of the following observations.

**Landlocked Status**

Often the borders of developing countries were created on the basis of which colonial power had control over them and the policies of that power and when it relinquished control. One of many determinants to a country’s level of development is whether it has access to the sea through navigable harbors and rivers, which is essential to be connected to the rest of the world, which in turn promotes commerce, investment, trade, and intellectual interaction. A landlocked country often has a difficult time in facilitating any kind of extensive exports and imports, at least without a friendly relationship with its neighbors. Africa alone has 15 landlocked countries, by far and away the most of any continent. Many of the world’s poorest countries are severely hindered by high transport costs because they are landlocked; are situated in high mountain ranges; or lack navigable rivers, long coastlines, or good natural harbors.

**Diseases**

Most of the tropics have ecological conditions that favor deadly diseases such as malaria, schistosomiasis, cholera, dengue fever, yellow fever, bilharzia, and dozens of others. Sub-Saharan Africa, in particular, has an ideal rainfall, temperature, and mosquito type that make it the global epicenter of malaria, one important factor in slowing Africa’s economic development. Almost every country that lies between the Tropic of Cancer and the Tropic of Capricorn is subject to the risk of one of these vector-borne diseases in the zones below 2,000 meters in elevation.
Malaria is a major international public health problem, causing 350 million to 500 million infections worldwide and approximately 1 million deaths annually. No vaccine is currently available to prevent malaria or dengue fever. Malaria is largely a tropical disease as it needs high temperatures, plenty of breeding sites, and mosquitoes that are attracted to humans. When children die in large numbers, parents compensate by having numerous babies. Too poor to invest in the education of all the children, the family might educate just one child, usually the elder son. Some development experts point out that a world map of low-income countries overlain with a map of malaria risk will reveal much commonality between the two.

Additionally, research from Kenya indicates that a malarial infection may lead to a higher incidence of HIV/AIDS cases; there may also be a reciprocal effect of HIV infection cases on incidences of malaria. HIV/AIDS is an endemic disease in sub-Saharan Africa, where in some countries up to 40% of the population is infected. This fact has devastated the social-cultural dynamics of many communities that now lack a whole middle-aged sector of society. In 2007 alone, the AIDS epidemic in Africa claimed the lives of an estimated 1.6 million people; more than 11 million children have been orphaned by AIDS. The extent of the AIDS crisis is only now becoming clear in many African countries, as increasing numbers of people with HIV are becoming ill. In the absence of massive prevention, treatment, and care efforts, the AIDS death toll in sub-Saharan Africa will continue to rise. This means that the impacts of the AIDS epidemic on these societies will be felt most strongly in the course of the next 10 years and beyond. Its social and economic consequences are already widely felt, not only in the health sector but also in education, industry, agriculture, transport, human resources, and the economy in general.

Plant diseases and pests are another type of obstacle that developing countries in the tropics are forced to confront. A case in point is the countries in Northern Africa such as Kenya, Somalia, Ethiopia, Sudan, and Eritrea, which are subject to annual infestations of locusts. A single swarm can cover 1,200 square kilometers and can contain between 40 and 80 million locusts per square kilometer. With each insect capable of eating its own body weight in vegetation each day, a swarm that size could consume 192,000 metric tons of vegetation each day.

### Drought

The distinction between the developed and developing worlds is often based on agricultural production. The production of food is contingent on the amount and availability of water for rain-fed and potentially irrigated agricultural land. The use of this potential is frequently not great in the developing world: Out of the 1,620,104,000 ha (hectares) in the developing world, only 207,432,000, or 3.9%, are considered as potentially irrigable. The reasons for this are variable. The terrain may be too steep, the soils may be inadequate, or the land may be arid or hyperarid. These conditions are not subject to change in spite of any number of external inputs.

Many of the least developed countries in the world are trapped in arid conditions with low agricultural productivity or are vulnerable to prolonged droughts. When there is no predictable pattern of annual rainfall amounts, it is very difficult for farmers to make long-term agricultural plans or to invest in capital improvements. Climatic studies in Africa indicate that long-term changes in rainfall have occurred in the semiarid and subhumid zones of West Africa. Rainfall during the 1968 to 1997 period was, on average, 15% to 40% lower than during the period 1931 to 1960. A similar but smaller change has occurred in semiarid and subhumid regions of Southern Africa. When there is such variability in the year-to-year mean amounts of rainfall, it is difficult to maintain a consistent level of food productivity.

The arable land per capita is also a major concern in many developing countries. The rapid population growth has meant that the arable land per capita has declined sharply in recent years. In 1961, for example, developing countries as a whole had an average of about one half of a hectare of arable land per person; by 1992, the amount had fallen to less than one fifth of a hectare. If current trends in population growth
and land use continue, in 2050 the amount of arable land will be just more than one-tenth of a hectare per person.

Economist Jeffrey Sachs argues that the biggest determinant to why some countries have escaped the ravages of poverty and others have not is food productivity. Soil fertility is the principal constraint for sustaining agricultural production in many developing countries. With rapid population growth, soil nutrient stocks are being mined due to a reduction in fallow lengths, cultivation of fragile lands, and limited use of inorganic or organic sources of fertility, lowering agricultural productivity and increasing poverty. For example, 19 of the 25 poorest countries in the world are located in Africa. Sachs says that unless the downward spiral of increasing poverty, declining agricultural productivity, and resource degradation is reversed, sub-Saharan Africa will continue to be the locus of poverty with the world highest proportion of poor people.

None of the basic resources required to expand food production—land, water, energy, and fertilizer—can now be considered abundant or inexpensive. In many developing countries, there has been a serious degradation of arable land. Population pressures have caused grave overexploitation of soils. Irrigation, overgrazing, and denudation of huge forest areas to obtain wood for fuel and to clear land for farming have further reduced the soil’s capacity to produce. Some of this land is marginal for farming, with soil and climatic conditions poorly suited for annual cropping. This is especially true where the more fertile lands are already crowded and the population spills over on to marginal land. Such land can only produce low yields and is more or less susceptible to degradation, depending on the quality of management.

International development experts who monitor progress toward a sustainable world say that soil erosion is slowly undermining the productivity of one third of the world’s cropland. Each year, the world’s farmers lose an estimated 24 billion metric tons of topsoil in excess of new soil formation. Deforestation is leading to increased rainfall runoff and crop-destroying floods. Between 1970 and 1990, the world lost nearly 200 million ha of tree cover and deserts expanded by some 120 million ha.

Conclusion

The factors that determine whether a country falls within the category of “developed” or “developing” are numerous. They include per capita income, population growth rates, education levels, trade policies, climate, governance, proximity to markets, outside investments, and effects of colonial history. However, the physical environment also plays into this mix and is one of the factors not often considered by economists and international development experts. It is, however, a salient issue that needs to be considered. The process can usually start with examining a map carefully and analyzing the spatial distribution of the physical features in relation to a given country.

An adverse geography poses problems that can be solved, typically through physical investments and good conservation management. But an adverse geography raises the costs of solving the problems of farming, transport, and health and thereby makes it much more likely that a country will be caught in a poverty trap. The diverse geography of the world also challenges development experts to think in a new paradigm. That is, what works in one country may not work in another. Solutions to the problems found in the “developing world” may not come out of what works in the North. Most likely, they will come out of each region with assistance (not direction) from the North.

William M. Van Lopik

See also Agriculture, Preindustrial; Colonialism; Decolonization; Dependency Theory; Development Theory; Disease, Geography of; Drought Risk and Hazard; Environment and Development; Fair Trade and Environmental Certification; Gross Domestic Product/Gross National Product; HIV/AIDS, Geography of; Informal Economy; International Monetary Fund; Malaria, Geography of; Malthusianism; Modernization Theory; Neocolonialism; Neoliberalism; Neo-Malthusianism; New International Division of Labor; Newly Industrializing Countries; Peasants, Peasantry; Plantations; Political Economy of Resources; Population and Land Degradation; Shifting Cultivation; Soil Degradation; Sustainable Development; Underdevelopment; World Bank; World Summit on Sustainable Development; World-Systems Theory
DEVELOPMENT THEORY

Development theories generally fall into one of two approaches. First, there are theories that provide explanations of the meanings, trajectories, and characteristics of socioeconomic progress and put forth policy frameworks to aid in the design of social and economic interventions to aid poor communities, regions, and less industrialized (i.e., developing) economies. Second, there are theories that critique mainstream ideas about development with the goal of demonstrating how economic and political powers/elites exploit the poor, working classes, and/or once-colonized peoples through ideologies and practices of development that ignore the real needs, voices, and knowledge of these individuals and their communities. In both cases, the primary focus is on communities and economies that were once colonized by European, American, or Asian powers (i.e., the developing world or Global South).

Origins

Development theory’s contemporary manifestations were inspired by three schools of thought about society and progress: classical political economy, historical idealism (Georg Wilhelm Friedrich Hegel) and historical materialism (Karl Marx), and sociological conceptualizations of societal evolution. Classical political economists (Thomas Robert Malthus, Adam Smith, David Ricardo, and John Stuart Mill) characterized societal progress (development) as being made possible through wealth creation, a divinely inspired process guided by the invisible hand of the market and improved or facilitated through international trade based on the comparative advantages of nations. For the classical scholars, development (i.e., progress) would occur if there was an adequate division of labor and effective economic circulation mechanisms (i.e., money, trade, and capital flows) to sustain wealth creation.

A second early inspiration came from Hegel and Marx’s theories of social evolution. For Hegel, humanity (but principally Europeans) evolved through a dialectical progression of ideas that would lead to the absolute idea, a point in history when humanity would be spiritually unified through the discovery of reason in its highest form. For Marx, a universal history, was possible but he rejected Hegel’s spiritual notions and argued instead that it was the historical-material conditions of life that would lead societies toward higher levels of consciousness, an ideal endpoint being the resolution of class struggles and capitalist exploitation through the emergence of a classless society.

The third important foundation emerged from sociological explanations for the evolution of modern capitalism. Max Weber, Thorstein Veblen, Michael Polanyi, and others sought to explain how the institutions of capitalism (e.g., property, money, firms, classes, or markets) emerged in Europe, the United States, and Japan and what their rise meant for social well-being. Modern development occurs when social rules, norms, and institutions evolve to guide individuals toward the rational decisions needed to sustain and support the capitalist system. Particularly significant is the notion that modern development involves an increasing separation or disembedding of economic institutions (i.e., markets, firms, property rights) from political ones. However, as Polanyi observed, this separation is at best only partial as the crises and inequalities inherent in capitalism inevitably force the state to interfere with the “natural” market system.

Further Readings

Growth and Modernization Theories

The mid 20th century was a revolutionary period for development as growth and modernization theories came of age. John Maynard Keynes’s theories on economic multipliers, full employment, and the role of the state in guiding the creation and distribution of prosperity inspired the birth of growth theory. His protégés Roy Harrod and Evsey Domar posited that it is through appropriate levels of growth that full employment (i.e., optimal prosperity or development) is sustained, and their models sparked an intense debate within economics that led to the Solow-Swan growth theory. Robert Solow and Robert Swan argued that growth is driven by the productivity increases that accompany modern industrialization (e.g., increasing capital intensivity, technological change, knowledge accumulation, and labor efficiency improvements). Productivity growth remains a central theme in development economics, and it is viewed by many as the critical mechanism for generating and sustaining economic progress.

Early growth theories were important influences on the new generation of modernization theorists such as Arthur Lewis, Simon Kuznets, Albert Hirschman, Gunnar Myrdal, and Walter Rostow. For Lewis, modernization—a transition from a traditional (i.e., subsistence and/or agrarian-based) to an industrialized economy—would occur as the unlimited supply of low-wage labor existing in “backward” countries shifted into commercial activities sustained by favorable terms of trade. If the state intervened appropriately, the resulting profits could be reinvested in projects to reduce poverty and improve the commercial sector continuously. Kuznets, like Lewis, viewed underdeveloped countries as being structurally different from modern societies. His most lasting contribution is the Kuznets curve, a model showing that increases in inequality are inevitable at the beginning of modernization but that inequality eventually decreases in the transition to an industrial economy. For Hirschman, geographic inequality or the polarization of modernity within countries was a necessary first stage toward widespread development. Myrdal too recognized the inevitability of such regional inequalities and argued that states must intervene aggressively to mitigate their negative consequences. In his infamous attack on Marxism, Rostow posited that while inequalities were inevitable during the modernization process, these would dissipate as a large middle (consumer) class emerged through economic growth, industrial diversification, and an effective liberal democracy. For Rostow, every underdeveloped, noncommunist society could modernize independently through five stages of growth that would effectively transform the economy from a traditional system to one driven by mass consumption and modern industries.

During the 1960s and 1970s, geographers such as Brian Berry, Lawrence Brown, Peter Gould, and Akin Mabogunje played an important role in development theory by analyzing, mapping, and theorizing about the spread of modernization within developing countries. These studies were inspired in large part by François Perroux’s regional growth pole theory, Walter Christaller’s ideas about central places, Hirschmann and Myrdal’s work on polarization and cumulative causation, and Torsten Hagerstrand and Everett Rogers’s studies of innovation diffusion. Common research themes included analyzing and mapping the spatial distribution of modernization in cities and regions, studying the demand and supply-side mechanisms that facilitated the diffusion of innovations, and developing spatial models to maximize the efficiency of transportation, communication, and other distribution systems.

Neoliberalism and the New Institutionalism

In the 1980s, neoliberalism became the dominant ideology worldwide shaping mainstream policies of institutions for development. Early neoliberal theorists believed that state-led import substitution industrialization programs had failed to deliver the promised levels of modernity and that freer markets and trade could more efficiently drive development. Inspired in part by the works of the economist Friedrich Hayek, neoliberal thinkers argued that market forces would be the most effective way to advance social progress and that states needed to allow the private sector to guide development. International trade—in as “free” a form as possible—is central to these theories, and developing countries were encouraged to
DEVELOPMENT THEORY

liberalize import markets and industrialize through export-led strategies. Citing the “miraculous” success of the export-oriented East Asian “tigers”—South Korea, Taiwan, Singapore, and Hong Kong—the Washington Consensus (the International Monetary Fund, the World Bank, and the U.S. Treasury) promoted an aggressive free trade agenda instituted in part through the establishment of the World Trade Organization (WTO).

Neoliberalism’s promoters quickly realized that achieving a transition from a state-led to a market-driven development process required a dramatic transformation of the institutions structuring and regulating economies. Inspired in large part by the work of scholars such as Ronald Coase, Oliver Williamson, Douglass North, and Robert Bates, the “new” institutional economics became an important theoretical influence on neoliberal (and other) development theories. Institutionalists believe that individuals and firms make rational choices in response to the cost and price signals transmitted to them through the markets, rules, norms, laws, and conventions (i.e., the institutions) structuring socioeconomic activity. Successful neoliberal institutions clarify and protect property rights and reduce the transaction costs individuals encounter when making consumption or production decisions (i.e., choices that create profits or an economic surplus). These theories, when coupled with neoliberal ideology, translated into development policies (e.g., structural adjustment programs) aimed at dramatically restructuring market, state, and civil society institutions to increase economic efficiency, raise productivity, and spur rapid integration into global markets.

Social capital is constituted by trust, social networks, norms of reciprocity, and levels of civic engagement, and it can facilitate progressive development if it is widely distributed throughout society and maintained, reproduced, and increased through appropriate state and nongovernmental interventions. Development projects can increase social capital by providing institutional mechanisms to connect a diversity of individuals, firms, markets, states, and communities in networks characterized by high levels of trust and reciprocity. In contrast to modernization and neoliberal approaches, these theories deemphasize material worlds and market forces and focus on how historically, socially, and geographically embedded networks, civic institutions, and cultures shape development processes and the ability of a society to create external ties to other places and regions.

Sustainable development theory is a second significant area of contemporary research. This field grew out of increasing concerns in the 1970s and 1980s about global environmental problems and the long-term prospects for development through modernization. With the publication of the Bruntland Report in 1987, the concept of sustainable development entered the mainstream, and the field has grown dramatically in scope and scale since that time. Sustainable development’s advocates believe that long-term human progress can only be achieved if economic desires and social needs are in balance with environmental limitations such that present-day development is not achieved at the expense of future generations’ abilities to do so. Central to these debates is the question of whether development through economic growth, driven by the production and consumption of natural resources, is possible indefinitely. Weak proponents of sustainable development believe that growth and sustainability are compatible, provided technological changes enable humans to find substitutes for declining or nonrenewable resource stocks. Strong proponents of sustainability believe that there are profound limits to growth and that long-run development will only be possible if humans create and sustain an ecologically and thermodynamically balanced economic system (i.e., a steady-state economy).

A third strand of contemporary thinking emerged in the context of postcolonial and
postdevelopment critiques of modernization and neoliberal theory. While there are differences between the postcolonial and postdevelopment approaches, there is general agreement in both communities that modern development theory has been essentialized, universalized, and violently exercised by colonial and Northern economic powers who view it as a necessary transition from one binary state to another (e.g., from traditional to modern; backward to progressive). These theorists argue that development, if there is such a thing, should be an emancipatory process through which the knowledge, voices, and livelihoods of the subaltern (people in the Global South who have been subjugated to colonialism and exploitation by the global economy) are understood, appreciated, and empowered through direct political action. To do so, development policy and theory must be decentered away from Western experts, neoclassical and neoliberal economics, and hegemonic multilateral institutions such as the World Bank and the International Monetary Fund. Once these powerful institutions are destabilized, it may be possible to transform development thinking and practice progressively through an understanding of the subjectivities, social movements, cultures, power relations, and knowledge situated in the postcolonial world (i.e., developing regions).

Critiques

Critiques of development theory have been significant and continuous since the 1950s. Central to these criticisms has been the concern that development theory and practice are not meant to help the poor or disempowered but are instead pursued by powerful states as a means of perpetuating exploitative modes of capitalist production, geopolitical relations, and meanings of “progress.” In broad terms, these criticisms have come from two primary perspectives—radical political economy and postcolonial and postdevelopment studies.

The radical political economy of the 1960s and 1970s drew heavily on classical Marxian theory and responded directly to modernization and growth theorists. Dependency theorists such as André Gunder Frank and Raúl Prebisch viewed modernization schemes as neocolonial mechanisms for sustaining unequal exchange relationships between poor and rich countries. Specifically, modernization and international trade were tools to encourage developing countries to export cheap raw materials while keeping them dependent on wealthy economies for more expensive manufactured goods and services. Because of this unequal exchange, the capitalist trading system linking metropole states (industrialized economies) and satellite economies (the developing world) would inevitably and perpetually siphon surplus value away from poor countries even if they established more advanced industrial sectors. Thus convergence between rich and poor, what Rostow argued in his stages-of-growth framework, was never really possible, and developing economies were better off focusing on strategies to promote autarky or self-reliance.

Dependency theory was extended by Immanuel Wallerstein through the development of world-systems theory in the early 1970s. World-systems theory posits that the world economy is organized into a tripartite system of power relations (core economies (most powerful), semiperipheral economies, and peripheral economies (weakest) that historically evolved through geopolitical and geo-economic conflicts between countries seeking to maintain or increase their global power. As such, and in contrast to Rostovian notions that internal factors alone determine a country’s ability to develop, development is determined in large part by a country’s ability to overcome the external structural constraints (e.g., unequal exchange relations) imposed on it by more powerful countries and global capitalism.

More recent radical critiques have drawn on Michel Foucault, Antonio Gramsci, and political ecology literatures to attack the strategies of neoliberal institutions and expose the social, environmental, and economic outcomes of their policies. These scholars have placed particular emphasis on the role of multilateral development institutions, neoliberal states, multinational corporations, and empowered development “experts” in perpetuating unequal power relations between haves and have-nots in the global economy. Such critiques, while they are not without flaws, have exposed significant weaknesses in neoliberal development theory and demonstrated why it has not led to a dramatic redistribution of power, wealth, and opportunity in the world economy.
Radical scholars have also drawn on political ecology theories to show how neoliberal development strategies are adversely transforming ecosystems and human-environment relations throughout the global South.

Beyond radical political economy, development theory has been critiqued through the lenses of postcolonial and postdevelopment studies. Postcolonial and postdevelopmental theorists have examined the historical and cultural production of the power relations driving development processes and detailed how modernist visions of progress objectify once-colonized peoples in the Global South (the subaltern), effectively silencing their voices in debates and discussions about what development is and how it might be achieved. As Paul Farmer argues, development interventions are often forms of structural violence against the institutions shaping the everyday realities, understandings, and rationalities of poor people, not liberating projects to empower and improve the lives of the subaltern. Moreover, because development experts lack sufficient reflexivity and humility about how projects are designed, funded, and implemented, their understandings of the “targets” of development programs (poor people, public health, etc.) are at best partial, and it is this partiality, hubris, and rigidity that doom many development interventions from the outset. Postcolonial and postdevelopment critiques have raised important concerns about how development is practiced and forced a critical rethinking of what development is, how it might best be achieved, and how the poor might take real control over the process.

James T. Murphy

See also Dependency Theory; Developing World; Diffusion; Economic Geography; Export-Led Development; Foreign Aid; Globalization; Growth Poles; Import Substitution Industrialization; Industrialization; International Monetary Fund; Modernization Theory; Neocolonialism; Neoliberalism; New International Division of Labor; Newly Industrializing Countries; Orientalism; Political Ecology; Political Economy; Population, Environment and Development; Postcolonialism; Poverty; Regional Economic Development; Rural Development; Spatial Inequality; Structural Adjustment; Subaltern Studies; Sustainable Development; Technological Change, Geography of; Underdevelopment; Uneven Development; World Bank; World-Systems Theory; World Trade Organization (WTO)

Further Readings

Diamond, Jared (1937– )

Jared Diamond is a well-known physiologist and evolutionary biologist. In addition to his research in these fields and in conservation ecology and anthropology, Diamond has published a variety of works on biogeography and human geography. He has been a professor of physiology in the medical school at the University of California at Los Angeles (UCLA) since 1986 and professor of geography in the department of geography at UCLA since 2001. Diamond also serves on the boards of directors of World Wildlife Fund and Conservation International.

Diamond has conducted more than two dozen expeditions since the 1960s to research biological diversity in New Guinea. Recently, he led National Geographic Society and World Wildlife Fund projects to survey the distribution and spatial characteristics of bird species in Papua New Guinea. Notable among his resulting publications is *The Birds of Northern Melanesia: Speciation, Ecology, and Biogeography*, coauthored with Ernst Mayr. This tome compiles and updates more than a century of taxonomic data and uses these data to address questions about oceanic avian speciation and, more generally, island biogeography. Diamond’s ongoing surveys in this region are part of a broader conservation effort to curb loss of biodiversity; he has also contributed to the design and implementation of conservation parks in the Southwest Pacific.

Diamond is best known for his studies of human geography, particularly his global-scale studies of historical, cultural, and environmental geography. These include attempts to answer major disciplinary questions in the fields of archaeology and anthropology, history, and sociology through geographical surveys of the human present and past. His 1997 tour-de-force work is *Guns, Germs, and Steel*, which examines the historical contingencies of human biological variability at a worldwide scale. Diamond uses biogeo graphical and environmental patterns to account for the cultural-historical trajectories of the New World versus Old World. The book won a Pulitzer Prize in 1998 and is hailed as an unparalleled, global reexamination of human geography spanning the past 14,000 years. In a similar vein, his 2005 book *Collapse: How Societies Choose to Fail or Succeed* examines the historical geography of societal disintegration through a wide survey of ancient and modern case studies. Here, he links social collapse to human decision making in the context of ecological and geographical patterns. Parallel to his work in New Guinea, *Collapse* ultimately addresses modern conditions and future prospects, outlining ecological management, environmental stewardship, and other social solutions to future global crises.

Diamond regularly contributes popular articles on topics ranging from evolutionary biology and the origins and spread of human disease to commentaries on current discoveries and research in *Nature* and other scientific journals. He has received numerous prestigious awards, most notably a MacArthur Foundation Fellowship (1985), the President’s National Medal of Science (1999), the Cosmos Prize (1998), and the Alfred Russell Wallace Award by the International Biogeography Society in 2007.

Scott Van Keuren

See also Biogeography; Environmental Determinism; Environmental Management; Environment and Development; Island Biogeography

Further Readings


Diaspora

Originally used to describe the scattering or dispersion of Jews from Palestine in the 6th century BC, the term *diaspora* has evolved into a widely used reference point for understanding human migration, settlement, and transnationalism. Since the late 20th century, the term *diaspora* has become a common label for any ethnic population living in places other than that considered their historic homeland and also refers to
collective practices that challenge the boundaries of the nation-state. The broad and growing interest in diasporas has helped forge diaspora studies as a distinct academic discipline; however, other disciplines, including geography, have also made important contributions to the subject.

Academic studies on diaspora center on several themes, including migration, loss, memory, settlement, and transnational experiences. Migration and settlement themes focus on the historical origins and experiences of diasporas as they develop as unique cultural and political communities outside “native soil.” These experiences of displacement and resettlement furthermore provide context for studies that examine the loss and remembering of cultural practices and meanings. Recently, the notion that diasporas are communities that characteristically suffer cultural loss has given way to the view that they form hybrid identities, mixtures of the cultures of their origin and the host countries. Furthermore, many current studies concentrate on how members of diasporas construct their own personal identities both in relation to their diaspora communities and as participants in the construction of the broader significance of their diaspora.

The theme of transnationalism has become especially important in contemporary dialogues about diasporas. The idea that diasporas involve complex interrelationships between people and places, and often across political and socio-cultural boundaries, predates the concept of transnationalism. Yet, with the recent interest in transnationalism as an aspect of globalization and contemporary migration, the transnational nature of diasporas has attracted wider interest. One reason for this renewed interest is that studies on the transnational practices undertaken by members of diasporas acknowledge the symbolic as well as material interchanges among places. In particular, the treatment of transnational ties in studies concerning diasporas often include emotional and psychological dimensions that are elsewhere often neglected.

One of the remarkable details about the diaspora concept is how quickly the usage of the term has expanded to represent many different migrant and ethnic groups. The extension of the term has been associated with increasing migration and cosmopolitanism, but some have also suggested that it reflects the broadening alienation of people from meaningful political representation within existing nation-states.

In addition to the recent expansion of usage of the term in academic circles, there has been a remarkable rise in the attention of policy makers. Since 2005, several global institutions, as well as a number of countries, have published reports and policies acknowledging diasporas as significant agents of change in their countries of origin. Geographers have been among the small number of social scientists who have examined the rise of these policies, which are sometimes referred to as “diaspora strategies.” The growing and far-reaching interest in diasporas, as social phenomenon, political constituency, and analytical concept, underscores the deeply political nature of this subject.

Alexander Malcolm Lovell

See also Ethnicity; Hybrid Geographies; Immigration; Migration; Transnationalism

Further Readings

Diastrophism or tectonism is the group of mechanisms by which large-scale portions of the Earth’s crust are deformed by natural processes. Such diastrophic processes lead to the formation of continents and ocean basins, mountain systems and rift valleys, and other related features by the chief mechanisms associated with the plate tectonic movements of the lithospheric plates of the planet. The study of diastrophism, or of tectonic processes, is the central unifying principle in much of modern-day geology and geophysics.

Five main kinds of diastrophic processes occur:

1. Orogeny, in which narrow mountain belts are formed
2. Epeirogeny, which is broad regional uplift or down warping of the crust
3. Isostasy, which is broad down warping or uplift related to loading and unloading of the crust
4. Eustasy, which is worldwide sea-level changes from addition or removal of water or through tectonic change in ocean basin volume
5. Igneous processes, which include intrusive ones, or plutonism, which are emplacements of molten magma at depth in the Earth, and extrusive or volcanic emanations of lava and pyroclastics on the surface.

The main driver of all diastrophic processes is the movement of the great tectonic plates, whereby new rock is created at the spreading centers and old rock subsides down subduction zones to be melted and partially recycled.

Mountain building orogenesis is the group of processes whereby plate tectonics drives up the great ranges as the moving lithospheric plates rift apart in some places and collide in other places. Thus there are five main kinds of orogenies:

1. Uplift and rifting of a continent above a spreading center, which produces moderate-sized mountain uplifts and volcanoes, much like those in East Africa
2. Island arc growth into a cordillera above a subduction zone at a continental margin, as happened with the ancient Appalachian mountains
3. Collision of a volcanic island arc with a continent, as happened in the Andes and elsewhere
4. Collision of two continents, as happened with India as it moved into Asia and pushed up the Himalayas
5. The tectonic aneurysm, as happened at both ends of the Himalaya, where great river systems such as the Indus and the Tsangpo-Brahmaputra were captured and diverted so that their ferocious denudation unloaded the crust so quickly that its pressure release at depth resulted in igneous and metamorphic effects and a mountain pop-up on the surface.

Epeirogeny in which the crust is broadly and slowly up-warped or down-warped without significant faulting, folding, metamorphism, or igneous activity and other tectonism is not well understood. It may reflect dynamic effects of motions in the mantle below the crust, phase transitions in the mantle that the crust accommodates, or be the result of variations in the white-hot and plastic aesthenosphere below the lithosphere as the surficial plates move laterally over it. In any case, areas such as the Colorado Plateau and much of Central North America that were once below sea level are now well above it, which reflects their epeirogenesis.

Isostasy, the balance in Earth’s crust, which is likened to icebergs floating in water, is the situation wherein any great load such as water or ice added or removed from the crust will result in up-warping or down-warping as a result. The great glacier masses on North America and the Scandinavian countries during the Pleistocene, for example, greatly depressed the crust, with the result that Hudson Bay and the Baltic Sea still represent remnants of that down-warping once the ice melted away. Compensatory isostatic uplift is slowly under way in both places to restore the crust to the higher level it was at once prior to the advent of the glaciation. Similarly, in the state of Utah, Great Salt Lake is the small remnant of the once-huge, Lake Superior–sized, Lake Bonneville of the Pleistocene. Prominent wave-cut beach lines that are now high and dry were established at the high stand of this lake. Those old beach lines around islands in the deeper center of Lake Bonneville are now much higher in elevation than
those on the lake’s periphery, which is a reflection of the greater isostatic rebound where the water was once much deeper.

Eustasy, or worldwide sea-level change, has occurred throughout geologic time as water from the sea has been locked up in glacier masses up on the continents or has melted and returned to the sea to raise the sea level. Similarly, as the continents have alternately rifted apart or drifted together and ocean basins have come and gone, the volume of ocean basins has also changed as subsurface volcanoes grew in them and displaced the water onto the land. Thus, at times in the past the seas have alternately covered the continents and withdrawn from them to expose their marine sediments in many places worldwide.

Igneous processes are a major part of diastrophism in that a significant part of the rocks on the Earth’s surface have once been molten magma intruded at depth or extruded as lavas and pyroclastic fragments at the surface. The plutonic igneous intrusions are classified in part on the basis of their size and their relations to the layers of the surrounding country rock into which they intrude. Thus the discordant intrusions that cut across the layers of the country rock are the tabular dikes and cone sheets, as well as the irregular small stocks and large batholiths, whereas the concordant plutons are tabular sills, mushroom-shaped laccoliths, and the bowl- or funnel-shaped lopoliths. The volcanic igneous extrusions are the surface manifestations of intrusive processes at depth and are classified on the basis of their size, the steepness of their sides, and the types of eruptive activity. Thus ash and cinder cones are composed of explosively ejected pyroclastic (fire broken) fragments that pile up steeply as relatively small accumulations. Shield volcanoes are large, broad areas where fairly fluid lava effusions have accumulated on gentle slopes. Lava domes are small, steep-sided volcanoes where viscous lavas have piled up quietly, although if pressure builds up beneath such lavas, they can explode violently. Composite volcanoes or stratovolcanoes are those that combine elements of the other three types so that mixtures of explosive and effusive eruptions combine to build large, steep-sided volcanoes.

John F. Shroder

See also Geothermal Features; Global Sea-Level Rise; Oceans; Plate Tectonics; Volcanoes

Further Readings

Difference is a measure by which individuals, societies, and even nations seek to distinguish themselves. It is a measure of separation (as being unlike someone) and distinctiveness. In its assertion it creates an “other”—those we are not. Forms of difference have been grouped into broad social categories such as class, gender, race or ethnicity, and sexuality. However, difference can be asserted using any criterion, such as language, nationality, birthplace, religion, ancestry, and profession. It can also be tied to particular places and operates across many scales. This entry examines the concept of difference and the many types of difference that geographers have studied. It then discusses the various theoretical approaches that have been used in this field and considers its significance for future research.

For geographers, difference is an important analytical concept. Its examination is used to assert that individuals’ distinctiveness matters: to their experience, and the constitution, of particular places; to their life opportunities; and to the functioning of communities and societies. Individuals, and particular groups, will experience particular social processes differently (such as the closure of a factory that might affect men and women in different ways), and these differences will occur unevenly over space. The geographies of difference can be understood by exploring what differences have been studied and how these
have been measured and analyzed. There are also a number of approaches to exploring difference and an ongoing debate within geography as to whether we need to move beyond difference as a field of enquiry.

Geographers use the study of difference to expose the generalizing assumptions of earlier geographical work—some of which failed to acknowledge the diverse experiences of certain groups, such as women or the disabled. Thus difference as a concept is used to highlight, and fill, the silences and gaps in geographical scholarship. It is also used to understand how particular groups suffer because of their difference—through oppression or exclusion—in other words, “geographies of exclusion.” This understanding of what divides us also enables us to explore how we might build connections across differences and how we might overcome tensions between us (such as racial or religious divides) and thus live together more harmoniously. In this sense geographies of difference can take on a particular political project by being part of the debate as to how to build better societies. However, there remains a problem between valorizing difference and the need to understand the commonalities of human existence. For example, if we give too much credence to difference we can undermine the need for universal principles of justice and the power of working together (e.g., through collective action).

Geographers have explored a multitude of differences. The most frequently studied categories of difference have been class, gender, race, and sexuality. However, recent work has explored previously absent categories, for example, “old age,” and there is a nascent field of children’s geographies focusing on childhood and youth. Geographies of disability have also been investigated, including for example, reconceptualizing what it is to be deaf. In addition, what is studied within these categories and conceived of as different changes over time. These changes are in response to new contexts that give rise to new tensions. For example, since the terrorist attacks in the United States in September 2001 and in London in July 2005, the concept of race has been complicated by the emerging importance of religion as a defining marker of difference. Moreover, geographers have sought to interpret difference more broadly to incorporate majority groups, investigating masculinity and “whiteness.” Thus, how we conceive of difference continues to evolve.

Theoretical and Methodological Approaches

The ways in which these categories have been explored and analyzed have changed markedly over the past few decades. There remains much debate as to the most appropriate theoretical and methodological ways of perceiving and measuring difference. While there may be a spatiality to the differences that can be mapped and charted quantitatively (such as migration or housing patterns by category), more often geographers have taken a qualitative approach to exploring difference. This approach has involved understanding the processes, relations, and experiences that divide people and that shape and affect places and spaces. Such work has sought to bring the specificity of difference into view, to identify processes of exclusion, prejudice, or, conversely, solidarities and the assertion of rights and to understand how such differences are created and maintained and their consequences. In other words, differences have been explored through the examination of categories, spatial patterns, relations, and processes.

This variety of approaches can be illustrated through the example of one category—gender. In the late 1970s, gender emerged in geographical scholarship as a category of social difference. At first gender was explored through the absence of female academics in geography and the absence of women as a valid topic of geographical research. Feminist geography emerged as a field that asserted the need to explore gender relations and inequality in work, housing, and everyday life. Next, gender as a construct began to be problematized. Geographers sought to examine how the understanding of gender was socially constructed and contested, and thus studies of femininity and masculinity emerged. Moreover, the intersections between gender and other social differences, such as sexuality, class, and race complicated gender as a discrete identity. Recently, overlap with research on the geographies of sexuality has encouraged more work to explore the embodiment of gender, gender as a negotiation of our bodily beings, and how gender is performed.
(expressed, practiced, and displayed) by different groups. From this very brief trajectory, it is clear that there are a number of alternative approaches to the exploration of gender as a form of difference and that these approaches can challenge each other.

It is possible to identify a number of approaches to understanding difference and, in particular, the geographies of difference. Difference is most easily defined as being in opposition to others around us. In this way it is relational as it only makes sense if we define how different we are in relation to others. This can be taken further to understand that the context (or place) in which we exist shapes how we understand those differences. Thus, people will identify commonalities and differences in part due to the place they are in (and those differences in turn shape how place is constructed). This notion becomes most obvious when we are displaced—we move to another place, and that movement makes us realize something about ourselves that we might not have particularly noticed previously. For example, we might not have considered our race if we grew up surrounded by those of similar race, but moving elsewhere to a place where our race is markedly different enables us to understand our race as a difference. Thus, place and, hence geography, is a key aspect of how we should understand difference, and difference is significantly informed by the place in which people exist.

If difference is relational, then it is also dynamic and fluid, rather than static and fixed. This is because if our understandings of difference can change according to the place we are in, they will also change over time or through encountering others in everyday life. So what it means to be a particular gender will change as we age, change location, change jobs, have children, and so on.

Difference can be understood as an assertion of a particular identity. This can happen in several different ways. People might unite around a common identity (such as being gay) despite differences (such as race, gender, and class). They might do this to assert political rights or in response to a personal need for expression and a sense of belonging. Identity may also be an outcome of several layers of difference on a broader scale, for example as a national identity. Geographers also increasingly understand that identity is not fixed; rather, we have multiple identities (such as being a woman, a mother, and an academic). Taking an identity approach enables us to understand how difference is a dynamic mix of choice, history, and place and is a result of power relations.

The notion of multiple identities illustrates the permeability of many categories of difference and how they intersect and overlap. Some scholars have argued that this degree of intersection is key to understanding how people experience difference. Thus, the negative implications of being different—oppression and exclusion—will be compounded by the interconnection of multiple forms of difference (such as being gay and black). Consequently, to fully comprehend difference we need a broad and complex examination of these intersections and power relations within society. Such an approach raises important questions about what broader processes might be occurring behind these categories of difference to which we should pay attention. The examination of power is a way to explore the relations of difference, but more than that it is potentially an approach that moves beyond difference and seeks instead to understand the unevenness of societies as a result of uneven power relations. An examination of power enables a nuanced exploration of why difference will manifest itself unevenly across social and spatial relations, and, thus perhaps most important, how the negative aspects of difference can be challenged.

Approaches to the geographies of difference continue to evolve. Just as difference came to be perceived as a social construction, there are those who challenge such assumptions or argue that we accept too easily what counts as sameness: We need to look beneath apparent similarity to understand more complex geographies. New theorizations of difference are emerging, such as the examination of fleeting encounters or temporary transient events as moments through which we can explore the possibilities of moving beyond difference or articulating new forms of difference that challenge existing conceptualizations. Perhaps examining the geographies of indifference is just as important an endeavor. For example, certain acts such as racism can be allowed to continue precisely because of a broader indifference to others. Scholars are also exploring the notion of hybridity, which rejects the traditional categories of difference such as gender and race and
instead focuses on the performative subjectivity of individuals and integration of seemingly irreconcilable elements.

Regardless of the new directions scholarship might take, the geographies of difference are important precisely because the categories of difference are so enduring. The categories represent the real divisions and inequalities that continue to shape people’s lives and the places they inhabit. Geographers continue to explore how differences emerge, why certain markers of difference “stick” more than others and matter more in some places than others. Some differences, such as class and gender, have longer historical stories that appear to constrain the possibility of their radical reinvention or change. Other forms of difference are embraced, reclaimed, and subverted as powerful identifiers and useful intersections of commonality. Whatever their histories, boundaries, or expression, the task remains for geographers to acknowledge difference and yet understand and assert what it is that we have in common.

### Jenny Pickerill

*See also* Cosmopolitanism; Disability, Geography of; Ethnicity; Ethnic Segregation; Feminist Geographies; Gays and Lesbians, Geography and of; Gender and Geography; Hybrid Geographies; Identity, Geography and; Nationalism; Orientalism; Race and Racism; Racial Segregation; Subaltern Studies

### Further Readings


---

**DIFFERENTIAL HEATING**

Land surfaces warm and cool much quicker than water bodies, leading to differential heating of the Earth’s surface. Water has a higher specific heat (capacity to hold energy) and circulates energy to greater depth than does land, and so water bodies such as oceans, seas, and large lakes can absorb more energy yet heat more slowly than land surfaces. Since energy is more evenly distributed throughout the water column, water bodies also release energy and cool more slowly than do land surfaces. Nearly three quarters of Earth’s surface is covered with water, and so oceans and other large water bodies represent immense reservoirs for energy storage and release, strongly influencing climates worldwide. For example, differential heating results in the development of a series of regional thermal high and low pressure systems that reverse positions during daily and seasonal cycles. Wind and precipitation patterns reflect these alternating pressure patterns, and areas where these patterns persist for several months may experience a monsoon climate.

Water bodies and land surfaces transfer energy to depth by distinct processes. The land surface is relatively immobile, and so energy is transferred by conduction from one particle to the next. Conduction is a relatively slow process, resulting in most energy on land being stored near the surface. Energy is rapidly transferred and stored to great depths by convection in water via turbulence. Water is also relatively transparent, while land surfaces are opaque, and incoming solar radiation (“insolation”) penetrates deep into the water column but is absorbed at the surface on land. With more energy stored near the surface, land has much greater daily and seasonal temperature fluctuations.

The specific heat of water is also much higher than those of typical land surfaces, meaning it requires more energy to heat water than to heat land by equivalent amounts. Because of the high specific heat of water relative to the land surface, the temperature change is less even though water bodies typically absorb and release more energy than does the land surface. During the daytime and summer, when insolation is greatest, water bodies absorb and store vast amounts of energy
DIFFERENTIAL HEATING

with only moderate increases in temperature. At night and during winter, a portion of this stored energy is released with only minor cooling of the water body.

During periods of maximum insolation, more energy is transferred to the atmosphere by evaporation over water bodies than over land surfaces. As water molecules change from a liquid to a gaseous phase, energy is used to break the molecular bonds attracting individual water particles, causing the water temperature to decrease as energy is transported to the atmosphere as latent heat. Therefore, during periods of high insolation, increased evaporation diminishes temperature increases of water bodies and further intensifies land-water temperature contrasts.

Differential heating results in temperature and precipitation patterns for coastal areas distinct from locations in continental interiors, particularly in the midlatitudes. Water heats and cools much slower than does land, and, consequently, coastal areas experience only moderate annual temperature variations. In addition, rainfall is typically more evenly distributed throughout the year in coastal areas because of the proximity to extensive moisture sources. In the continental interior, temperatures fluctuate considerably in response to seasonal variations in insolation rates, and most precipitation is delivered as convective summer storms. For example, San Francisco, California, along the west coast of the United States, and Kansas City, Missouri, in the Central United States, are at similar latitudes and have similar mean annual temperatures. However, Kansas City has a humid continental climate, with cold, dry winters and hot summers with increased precipitation, while San Francisco experiences a relatively Mediterranean climate, with similar temperatures year-round, moist winters, and dry summers.

The pattern and degree of differential heating is directly related to daily and seasonal insolation variations. Wind and precipitation patterns along coastlines are particularly responsive to these variations. During the daytime the land surface is heated by insolation, which warms the overlying air, causing it to expand and rise to form a thermal low-pressure system. As a result, a sea breeze develops because the warm air rising over the land surface is replaced by cooler, moister air originating over the ocean. A sea breeze usually begins in the morning as soon as temperature and pressure differences are great enough to cause a horizontal airflow, and winds reach their maximum strength in the late afternoon, when the differences are greatest. Sea breezes cease shortly after sundown, when the land surface cools to a temperature close to that of the ocean. In the night the land surface cools quicker and the wind pattern reverses as a thermal low develops over the now warmer ocean, and a land breeze forms, with winds blowing from the land to the sea. However, land breezes are not as strong as sea breezes since overnight temperature contrasts between land and water are typically less. In the summer, due to the insolation being the maximum, the temperature differences between land and water are greatest and land and sea breezes are strongest; during winter, the temperature differences are not as great and the winds are not as strong. The precipitation patterns are similar to the wind patterns, with the most precipitation delivered in late afternoons, slightly lagging the time of the most intense sea breezes. Also, more precipitation is delivered during the summer, when sea breezes are strongest. Skies typically clear overnight, when a land breeze develops, and less precipitation is delivered in winter, when sea breezes are weakest.

If differential heating persists for several months, such as throughout winter and summer, prevailing wind and precipitation patterns can develop that result in a monsoon climate. Monsoon winds and rains occur over large land masses and vary seasonally rather than daily. A semipermanent thermal low-pressure system develops over land during the summer draws in cool, moist air from the ocean and causes prolonged intense precipitation. As the land cools in winter, cool, dry airflows to the ocean, resulting in persistent clear, dry weather. The Indian subcontinent and southeast Asia, with clear skies in winter and intense summer precipitation from moisture-rich air originating over the Indian Ocean, provide some of the most dramatic examples of monsoon climates driven by differential heating. Monsoon weather patterns affect almost all tropical and subtropical regions, including parts of North and South America, sub-Saharan Africa, and Australia.

Mark W. Bowen
The term *differential vulnerabilities to hazards* refers to the differences in impacts of various types of hazards on different segments of society, which are often related to class, gender, ability and disability, race, age, and geographic location. For example, the greatest users of electricity are the wealthy, but the places of extraction and power plants are not normally near them. The same is true of consumption of goods as the wealthy use the most, but landfills and dumps are rarely near them, whether one considers the Cherry Island landfill in Wilmington, Delaware, in the United States or the pile of smoldering trash at Smokey Mountain in Manila, Philippines, that has collapsed and killed the poor people living off its meager resources. Such differential vulnerabilities are important to geographers in part because of the way that they illuminate factors affecting society’s exposures and consideration of ways to mitigate hazards and in part because such issues relate back to human rights and questions of good governance.

Hazards fall broadly into the categories of immediate and creeping hazards. Creeping hazards, such as desertification, occur so slowly that their onset is not recognized, whereas an immediate hazard is very apparent and often elicits a significant response. Vulnerability to hazards is determined by a combination of human and biophysical systems. There are many ways to dissect this confluence of factors. One way to examine vulnerability is to contrast wealthy-industrialized and less wealthy geographic and technological contexts.

### Less Wealthy Versus Wealthy Societies

In the less wealthy parts of the globe people tend to be relatively directly dependent on the natural environment. Their resilience is mostly locally determined, and buffers against hazards are thin in terms of time, as well as economic and material resources. For example, when powerful storms such as Typhoon Chataan strike a less wealthy and nonindustrial place such as the Federated States of Micronesia (FSM), there is typically less warning, less locally based technology for emergency response, and less wealth to support nutrition and recovery. Consequently, people are caught relatively unprepared, and often their capacity to feed themselves and recover quickly is weak. It should be noted, however, that such people do at times develop their own low-technology systems to aid in coping with hazards such as drought (see photo).

In wealthier and industrialized parts of the world, people generally have more capacity to prepare for, and respond to, what are commonly termed major natural hazards, such as earthquakes and strong storms. They have greater infrastructure, and their wealth allows both recovery and massive importation of food from outside the local area. However, such places also create their own hazards through industrial processes and technologies.

An example of a creeping hazard is California’s preparations to manage the water supply during drought (climate change) through reservoir infrastructure. However, the state quietly and accidentally polluted significant portions of its groundwater supplies in the 1990s and early 2000s though the introduction of chemicals such as MTBE (methyl tertiary butyl ether) into the water supply, which was added to gasoline to reduce air pollution!
Circular breadfruit pit layered with stones on the island of Fefan, FSM, which can hold breadfruit for around a decade to avoid starvation

Source: Dr. William James Smith Jr.

A different example can be pointed to in New Orleans, Louisiana, in the United States, where Hurricane Katrina became the costliest hurricane in U.S. history during the last week of August 2005. The faulty levee systems of New Orleans created a sense of security, but this technology and the policies connected to it led to populations living in areas that flooded terribly, and many, especially African Americans, the disabled, and the sick, were trapped with inadequate paths for escape. Hurricane Katrina killed more than 1,500 people in Louisiana alone. The estimated population of New Orleans in 2000 was 484,674. According to the Rand Corporation, as of September 2008 its estimated population was only 272,000.

Contrast this with the tsunami of December 26, 2004, which struck from southeast Asia to Africa, hitting many poor communities. In this case it was the lack of technology and no or poor warning systems that led to a terrible toll on life. Some small atolls in places such as off the coast of India were endangered by their location and island height of perhaps 2.3 meters and were physically reshaped by columns of waves that passed over them. Overall, this tsunami killed 230,000 people, according to the United Nations.

Several risks associated with fossil fuels create differences in vulnerability to hazards between wealthy and less wealthy countries. The vulnerabilities of the environment, fuel prices, and fuel supply risks are relatively severe and difficult to overcome in less wealthy countries due to their capital and technology constraints. Air emissions from fossil fuel consumption pose serious, and
sometimes extreme and immediate, threats to both public health and natural ecosystems. Countries are increasingly exposed to fuel price volatility because most new electricity generation in recent years has been gas fired, and oil demand has kept increasing due to greater vehicle miles traveled and increasing car ownership (fuel price risk). Many less wealthy countries are heavily reliant on imports to meet current oil and natural gas demands. As a result, they are vulnerable to disruptions in international trade, as was dramatically illustrated by the oil crisis of the 1970s (fuel supply risk).

Natural Versus Technological Hazards

It is highly debatable whether a stark line can be drawn between natural and technological hazards and human adaptation. When traditionally defined, natural hazards are easily categorized by the layperson, and technological hazards can be seen as hazards such as automobile crashes and nuclear plant accidents. However, in an era of global climate change due to fossil fuel consumption, the lines between natural and technological hazards are blurred. Researching and adapting to such hazards requires attention to human and biophysical systems, including technology studies. If indeed systems for economic growth and high technologies are often developed to serve the interests of the relatively rich, then differences in vulnerability are likely to be systematic.

Environmental Justice

Differences in vulnerability to hazards represent the root causes of environmental injustice. If all persons have environmental rights, then not addressing the varying impacts of hazards while taking into account class, age, ability and disability, gender, or geography represents an unjust response in terms of hazard mitigation. Differences in vulnerability to hazards by subgroup can occur for many reasons, including systems for managing natural resources and food and water systems in particular. For example, women serve as the primary water and food collectors in much of the world. If contamination occurs in the fields where they work, then they will be affected to a greater degree than men.

Other important vulnerability distinctions include the relatively high susceptibility to infectious disease of children and the elderly due to weak immune systems. Also worth noting is the enhanced vulnerability of the disabled when it comes to water and sanitation and, as previously noted, in times when evacuation is required. Yet another example is the relatively greater exposure of non-English-speaking groups to contaminated fish due to the lack of signs that they can understand to warn them not to consume fish at particular locations. A stark global inequity in vulnerability to hazards occurs to the small island countries of the world living just above sea level, which are remarkably vulnerable to sea-level rise, while emitting a negligible amount of greenhouse gases. A similar argument can be made in favor of the less wealthy countries that emit significant levels of greenhouse gases but have forest or land “sinks” that offset their emissions to an extent that many industrialized countries do not. Clearly, the examination of hazards requires geotechnical capacity. However, without blending human and geotechnical knowledge, it is not possible to effectively evaluate differences in vulnerability.

William James Smith Jr. and Young-Doo Wang

See also Coastal Hazards; Drought Risk and Hazard; Environmental Justice; Environmental Rights; Floods; Global Sea-Level Rise; Hurricane Katrina; Natural Hazards and Risk Analysis; Tsunami of 2004, Indian Ocean; Vulnerability, Risks, and Hazards

Further Readings


DIFFUSION

Although diffusion research is a central part of modern social science, it has a long and sometimes contentious history, with intellectual roots in anthropology and archaeology as well as in cultural geography. Diffusion research studies both diffusion—how innovations move—and innovation itself, which is the creation of new technologies that allow humans to control their environment better, whether such technologies are artifacts (hardware) or simply ways to program how work is done (software). Its third concern is adoption—how different societies receive, reject, and modify innovations. Such studies are common not only in geography but also in agriculture, education, rural sociology, history, and business studies. Understanding the processes by which technologies are transferred—innovation, diffusion, and adoption—is highly desirable in policy development and management studies. Within geography, there have been two main “schools” of diffusion studies. The Berkeley School of cultural geography, which emerged at the University of California in the 1920s, focused mainly on early human technologies such as agriculture. The Lund School, more theoretical and mathematical, emerged at Sweden’s Lund University in the 1960s. In 2003, a new journal, Comparative Technology Transfer and Society, appeared, covering diffusion studies across a variety of disciplines but with a heavy interest in case studies of technology transfer.

The Nature of Culture

The origins of diffusion research lie in 19th-century attempts to understand the nature of culture. Did human culture traits evolve in a manner analogous to biological evolution or did they appear rarely and diffuse out from a very limited number of innovation centers? Two extreme schools of thought developed. An evolutionary position that postulated the psychic unity of humanity concluded that all human beings were innately and equally innovative and that innovation was either continuous or triggered by relatively exogenous variables, such as population pressure. Driven also by notions that contact between human groups in preindustrial societies was minimal, this position led to the notion that most invention was independent and diffusion was either nonexistent or minimal. This has been described as the “utopian” theory of culture change.

At the other extreme, the second position held that there were a limited number of innovation centers or culture hearths from which key innovations such as agriculture diffused. Some argued for only one. This is usually referred to as the idea of Kulturkreis—“culture circles” of diffusion outward from the innovation center. Much early mapping of culture traits was an attempt to trace them back to a definable innovation center. For example, the remarkable data collected by the Soviet botanist Nikolai Vavilov in the 1920s and 1930s indicated that there were only eight centers of agricultural innovation. The implications drawn from the fact of the limited number of centers of innovation were that humans were essentially un inventive, that innovation was rare, that contagious diffusion—literally catching an innovation in a manner akin to catching a contagious disease—occurred easily even among preindustrial societies, and that people readily accepted new ideas. These conclusions form the basis of the theory of cultural change based on epidemic or extreme diffusion.
Eurocentric and Other-Centric Diffusionism

Epidemic or extreme diffusionism has been accused of Eurocentrism. Such diffusionism contributed to a powerfully held and generally unquestioned belief that European societies were inherently more, perhaps even uniquely, innovative and that the way to understand world history for the past 1,000 years or so was as the result of the diffusion of superior ideas and technologies out of Europe. One of the clearest rejections of this idea, although it drifts toward Sinocentrism in consequence, came from the British scholar Joseph Needham, whose seven-volume work covers the massive number of Chinese scientific and technological innovations over time. This work clearly demonstrates that China was the source of many innovations we consider European. Additional examples of other-centric forms of diffusionism can be found.

The Eurocentrism argument implies that non-European folk were too unschooled or naive to adopt European ideas and technologies. Diffusionism appeared at the high point of European imperialism, with Europe convinced that its machine technology was vastly superior to any other way of doing things. However, most of the colonial administrators of the great age of European imperialism were derived from the gentries and petite bourgeoisies that had seen their position in European society seriously eroded by industrialization. They were not about to re-create in their imperial possessions the same destructive forces that had damaged their social standing at home. Equally, indigenous elites in colonized societies had no vested interest in adopting the technologies that had already so damaged their societies.

Diffusion as a Theme in Geography

The Berkeley School of cultural geography came to a set of conclusions regarding culture change different from those of either the utopians or the epidemic diffusionists, although some Berkeley-trained geographers veered toward the latter position. The Berkeley School, led intellectually by Carl Sauer, separated origins from dispersals, provenance from process. Process was central since the Berkeley School embraced a cultural historical geography. Berkeley geographers did not follow the idealist explanations of the Kulturkreis School, nor did they easily fall into the trap of Eurocentrism. The Berkeley School blended idealist, environmental, and social-structural explanations to come up with relatively sophisticated depictions of provenance. But innovation was seen as rare, and the result was that the process of diffusion was much more easily explained and modeled. Heavily influenced by the Berkeley School, British geographer Peter Hall has returned to provenance issues, focusing on the concentration of particular innovations in particular cities at particular historical moments driven by particular historical processes.

The Soviet economist Nikolai Kondratiev observed that there appeared to be half-century-long wave rhythms in the global economy, accelerations and decelerations of the underlying growth rates of prices. The Austrian economist Joseph Schumpeter suggested that the root cause of long cycles lay in innovations in technology that drove gales of creative destruction followed by the flight of investment capital from less productive arenas into more productive ones. Geographers have contributed to the study of such long cycles in the global economy and have argued that innovation and diffusion can be seen as the very motor of capitalist development.

The main focus of diffusion research tends to be pragmatic, in part concerned with how societies innovate and in part concerned with how such innovations move around and are rejected, adopted, and modified. The unique contributions of geography to studies of diffusion have been to resist extreme diffusionism, concentrate on innovation as a process, and, in the case of the Lund School, develop highly theoretical and mathematical techniques to model and predict the spatial diffusion of innovations over time. This has also led to studies of the spatial structure of communication, in particular how it has been mediated by the life courses of individuals and families over time.

Diffusion as a Theme in History

Historians studying major themes such as the rise of Western civilization have emphasized not only the interchange of technologies and ideas among
the major global cultures but also the tendency of cultures to “forget” their borrowings from other cultures. Business historians and historians of technology have made clear over the past 40 years or so that diffusion is commonplace in the development of the industrial world. But it has always been easier and cheaper to copy than to innovate, and it seems to the uninformed that such a process would be simple. The existence of informal communications networks, such as networks of letter writers, has been shown by literary historians to be responsible for much “under-the-radar” diffusion. The supposed simultaneous development of ideas and technologies in different parts of the world has often resulted from the exchange of ideas through informal networks.

Technology Transfer and Development

No less a person than Karl Marx believed that the transfer of British railroad technology to India would result in the rapid and easy transfer of vital other mid-Victorian technologies such as machine textile production, the consequence of which would likely be British industrial decline. Similar arguments have been advanced by modern observers with regard to the likely shift of the information economy, again to India because of its common use of English, the main global business language, as a second language. Other countries, first Japan and now China, but also others, have been suggested as the likely heirs to Western production of goods.

Such claims usually ignore the problems of technology transfer. American attempts to acquire British textile machine technology in the aftermath of the American Revolution show how difficult such transfers are. The problems lay not in transferring (smuggling) machines out of Britain but in establishing in America the complex infrastructure of knowledge, machine tool making, and skilled labor. Case studies show that transfers of complex technologies are never easy because of the likely disparities between the receiving and sending cultures. This is even true of almost equally sophisticated societies: The transfer of organic chemistry from Wilhelmine Germany to America during World War I was long, arduous, and often flawed, despite the huge pressure of the war to increase explosives production.

Technology Transfer in the Automobile Industry

The increased importance given to patent infringement and copying in today’s world means there is a vested interest in both concealing and revealing such copying. No company readily admits that its product is legally copied from a competitor, let alone that it copies illegally. Some of the best examples appear in the automobile industry and come from Japan but have to be extracted from a complex literature in business history, enthusiast magazines, and books. An example of such technology transfer is the development of the Japanese Datsun automobile. The Japanese company DAT Motorcar, needing to market a small car in the early 1930s, simply copied the Rosengart, a version of the British Austin Seven. The Japanese intended to market the car, initially named Datsun and later renamed Datsun, in Australia as well as Asia. Charged by Austin with patent infringement, DAT bought a license in 1934. After World War II, DAT Motorcar renewed the license, and their Datsun products displayed a strong Austin heritage. This status changed with the development of the Datsun 240Z, which had German origins both in its engine, modeled after the Mercedes, and in its body design. DAT Motors was never very willing to admit any of this. Only after much legal prompting did the company recognize that Albrecht Goertz, the designer of the BMW 507 and Porsche 911 sports cars, was largely responsible for the body design for the 240Z.

Conclusion

The clear prevalence of diffusion in industrial societies makes it hard to sustain utopian theories of cultural change without running foul of the reasonable assumption that some sort of equivalent to the law of uniformitarianism has to operate in the cultural realm; it makes little sense to argue that the processes observably at work today are somehow radically different from those at work in the past.

It makes sense to distinguish innovation from diffusion. It is equally clear that theories of contagious diffusion are idealistic and almost never operate in the real world. Barriers to diffusion are often high, in particular in the cultural realm, and
so it makes sense to consider adoption environments. Knowing what makes a group able and willing to accept an idea or technology is as critical as knowing what makes them reject it: At a popular level, the latter is summed up as the NIH (“not invented here”) syndrome. Not all messages sent will be received: An important component of the adoption environment is thus the prestige of the sender, what geographer Philip Wagner refers to as *geltung*. Understanding how technology transfer sometimes fails is important. Finally, innovations simply must fit into preexisting structures of time and space, which implies the displacement of previous innovations, which can be described as de-novation.

*Peter J. Hugill*

**Further Readings**


**DIGITAL DIVIDE**

By now, digital reality and everyday life for many people have become so thoroughly fused that it is difficult, if not impossible, to disentangle them. In this context, simple dichotomies such as “offline” and “online” fail to do justice to the diverse ways in which the “real” and virtual worlds are interpenetrated for hundreds of millions. Yet for many others—the familiar litany of the poor, the undereducated, ethnic minorities, and the socially marginalized—the Internet remains a distant, ambiguous world. Denied regular access to cyberspace by the inability to purchase a personal computer, the lack of technical skills necessary to log on, or public policies that assume their needs will be magically addressed by the market, the information have-nots living in the economically advanced world are deprived of many of the essential skills necessary for a successful or convenient life. While those with regular and reliable access to the Internet drown in a surplus of information—much of it superfluous, irrelevant, or unnecessary—those with limited access have difficulty comprehending the opportunities it offers, the savings in time and money it allows, the sheer convenience and entertainment value it provides, and the ability to acquire data from it, from bus schedules to recipes to global news.

As the uses and applications of the Internet have multiplied, the costs sustained by those denied access rise accordingly. At precisely the historical moment that contemporary capitalism has come to rely on digital technologies to an unprecedented extent, large pools of the economically disenfranchised are shut off from cyberspace. As the Internet erodes the monopolistic roles once played by the telephone and television, and as the upgrading of required skill levels steadily renders information technology skills necessary even for lower-wage service jobs, lack of access to cyberspace becomes increasingly detrimental to social mobility. Indeed, those excluded from the Internet may be more vulnerable than ever before to social forces they do not and often cannot perceive.

**The Global Digital Divide**

In 2009, roughly 1.7 billion people, or 26% of the planet, used the Internet on a regular basis. The United States continues its long-standing
position as one of the world’s societies with abundant access to the Internet. Inequalities in access to the Internet internationally reflect the long-standing bifurcation between the First and Third Worlds. While virtually no country is utterly without Internet access (although portions of Africa come close), the variations among nations in relative accessibility are huge. Outside the global core, the vast bulk of the world’s people, particularly those in the Third World, have little to no access, a reflection of centuries of colonial occupation and their modern institutional legacies. To speak of the Internet as liberatory in impoverished social contexts such as Mozambique or Bolivia, with high illiteracy rates and few telephones, is absurd. What is more, within such countries networks are invariably concentrated within cities, whereas the plurality, and often the majority, of the population lives in rural areas. With slow connections and out-of-date telephone systems, graphical information—which uses much more bandwidth than text—is virtually out of the question (see Figure 1).

Global access to the Internet is deeply conditioned by the density, reliability, and affordability of national telephone systems, which form the heart of the architecture of cyberspace. For this reason, the distribution of Internet hosts also mirrors the enduring legacy of the superpower bifurcation during the Cold War: Soviet-backed regimes distrusted the telephone, which allows two-way communication, and preferred television, which allows only one-way flows of information. Most Internet communications occurs along fiber optic lines leased from telecommunications companies (which carry 80% of international communications), many of which are state regulated, in contrast with the largely unregulated state of the Internet itself. Prices for access vary by the duration of the phone call, the distance, and the degree of monopoly: In nations with telecommunications monopolies, prices are much higher than in those with deregulated systems. The global move toward deregulation in telecommunications, which started with the breakup of AT&T and privatization of British Telecommunications and has spread rapidly in many countries, will likely lead to more use-based pricing and fewer cross-subsidies (e.g., between commercial and residential users), a trend that will likely make access to cyberspace less affordable to low-income users worldwide.

The constraints on Internet access are not only economic but also political, given that the

Figure 1 Internet penetration rates internationally, 2009. International patterns of the digital divide mirror broader economic differentials among countries.
Source: Author.
electronic dissemination of knowledge can challenge established relations of power. Many governments have come to fear the Internet for its emancipatory capabilities. The Chinese government, for example, stung by students’ use of fax and e-mail during the 1989 Tiananmen Square massacre, limits access to Internet nodes. Private satellite dishes are still illegal in China and Malaysia. Singapore, no model of democracy, censors electronic information with a national standards review board. In 1996, the Guatemalan government, relying on the state-owned telecommunications company, Guatel, made private satellite or telecommunications links to the Internet illegal.

### Digital Divide in the United States

Although Internet penetration rates in the United States (74.9% in 2009) are not as high as in Scandinavian nations, they remain higher than those in many other urbanized, industrialized countries, and Americans as a whole still constitute the largest and most influential national bloc of Internet users in the planet. Several factors have conspired to accelerate Internet usage in the United States among and within different social groups. Of these, the continued decline in the price of personal computers (PCs) looms large. Following Moore’s Law, which speculates that the cost of computers falls by half every 1½ years, PCs have become ubiquitous across the United States. Indeed, relatively fast, low-end machines are readily available for less than $600 in numerous retail outlets. With 574 PCs per 1,000 people in 2005, the United States stands second only to San Marino in terms of availability of this device. Almost 80% of Americans use a PC (although not necessarily the Internet) once or more per week either at work or at home, the vast bulk of the PCs being networked. Because the value of a network rises in proportion to the square of the number of users, the Internet and the PC made each other increasingly powerful and attractive. Simultaneously, the rise in user-friendly graphics interfaces such as Netscape greatly facilitated Internet access for the parts of the population lacking in sophisticated computer skills. Moreover, as the number of applications of the Internet has grown, the hours of usage have increased steadily to more than 9 per week. The rise in PC ownership has been central to those who argue that the digital divide will disappear on its own accord.

Throughout the 1995 to 2009 period, the growth in Internet use among various sociodemographic groups was rapid, often spectacular. Average Internet penetration rates—including access at home, work, or school—more than quadrupled, from 14% to 74.9%; by 2009, 253 million Americans were using the Internet regularly. Thus, the innovation, the most rapidly diffused technology in world history, went from a tool or toy of a minority to an essential implement used by the vast majority. Every social group, as differentiated by age, gender, race/ethnicity, educational level, or household income, experienced marked gains. Thus, to the extent that the digital divide persists in the United States (and other economically advanced countries), it must be understood within the context of this sustained and rapid increase in the number of users and proportion of the population (see Table 1).

This growth, however, did not occur at identical rates among all social categories. Take, for instance, age. The young (i.e., below 30 years of age) steadily exhibited the highest Internet penetration rates, reaching 83% in 2006. For many children who grow up surrounded by digital technologies, the Internet is hardly mysterious. In contrast, in both benchmark years, the elderly experienced the lowest rates of Internet usage (a mere 2% in 1995 vs. 33% in 2006), as well as the slowest rate of increase in users. Many elderly people find new technologies to be difficult or intimidating, do not appreciate the potential benefits, and are easily frustrated by their lack of technical skills. The digital divide, therefore, is closely wrapped up with generational differences.

Notably, gender differences in Internet usage, which included an eight-percentage-point lead among men in 1995, declined steadily throughout this period, so that by 2006 it declined to a relatively minor two percentage points. Despite its popular reputation as an exclusive haven of masculinity, the Internet has in fact been harnessed by increasing numbers of women. Gender differentials in access reflect both the lower socioeconomic status of women relative to men as well as sexist cultural attitudes toward science and technology. The declining gender gap speaks of the increasing
familiarity with digital technologies among many women, particularly the young and well educated, who are often employed in producer services in which minimal computer skills are an essential prerequisite. Moreover, enrollment rates in American universities for women have consistently surpassed those for men, indicating that the future gendered digital divide will become smaller yet, if not disappear altogether.

The dimension of the U.S. digital divide that has drawn the most serious scrutiny concerns racial or ethnic differences. Given the profound inequalities in U.S. society in terms of income, educational opportunities, and employment that exist between whites and ethnic minorities, it is not surprising that this gap is manifested in terms of access to cyberspace, that is, much of the racial ravine in digital access is due to income discrepancies. In 1997, for example, white Internet usage rates were more than double that of Latinos/Hispanics (37.7% vs. 16.6%) and roughly double that of blacks or African Americans (19.0%). In 2006, Internet access rates for whites remained well above those for minorities or the national average. There are signs, however, that this dimension of the digital divide is slowly, if hesitantly, diminishing. Today, the majority of ethnic minorities uses the Internet, and the relative difference between them and the white population has declined. There are important differences within minority populations, however. Among African Americans, Internet usage tends to be concentrated among the young and the college educated, particularly women. Likewise, the Latino population is far from heterogeneous, and significant discrepancies in Internet access and usage remain among various subgroups; usage rates tend to be much higher among bilingual Latinos than among those who speak only Spanish. Indeed, among English-dominant Latinos, Internet usage rates are identical to that of whites. Generally, Mexican Americans and those with origins in Central or South America had lower rates of access than did Cuban Americans or Puerto Ricans. In short, while racial or ethnic discrepancies in Internet access and usage remain, all groups have experienced significant growth in use, and the relative differences between them have declined.

Persistently underlying the digital divide in the United States are vast socioeconomic differences, particularly education and household income, which effectively serve as markers of class. Although populations at all four broad educational levels (less than high school, high school graduate, some college, college graduate) exhibited gains in Internet access, profound differences remain. Among college-educated Americans, Internet usage is almost universal (91%); those with a high school education or less witnessed a growth in usership from 2% in 1995 to 35% in 2006. Educational level, therefore, is a prime predictor of who is online and who is not. Similarly, income remains a useful measure of who

<table>
<thead>
<tr>
<th>Percent Online in</th>
<th>2006</th>
<th>1995</th>
<th>Percentage of Growth</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18–29</td>
<td>83</td>
<td>21</td>
<td>62</td>
</tr>
<tr>
<td>30–49</td>
<td>82</td>
<td>18</td>
<td>64</td>
</tr>
<tr>
<td>50–64</td>
<td>70</td>
<td>9</td>
<td>61</td>
</tr>
<tr>
<td>65+</td>
<td>33</td>
<td>2</td>
<td>31</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>70</td>
<td>14</td>
<td>56</td>
</tr>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Men</td>
<td>71</td>
<td>18</td>
<td>53</td>
</tr>
<tr>
<td>Women</td>
<td>69</td>
<td>10</td>
<td>59</td>
</tr>
<tr>
<td><strong>Race/ethnicity</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>72</td>
<td>14</td>
<td>58</td>
</tr>
<tr>
<td>Black</td>
<td>58</td>
<td>11</td>
<td>47</td>
</tr>
<tr>
<td>Latino/Hispanic</td>
<td>69</td>
<td>21</td>
<td>48</td>
</tr>
<tr>
<td><strong>Education</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than high school</td>
<td>36</td>
<td>2</td>
<td>34</td>
</tr>
<tr>
<td>High school graduate</td>
<td>59</td>
<td>8</td>
<td>51</td>
</tr>
<tr>
<td>Some college</td>
<td>84</td>
<td>20</td>
<td>64</td>
</tr>
<tr>
<td>College graduate</td>
<td>91</td>
<td>29</td>
<td>62</td>
</tr>
<tr>
<td><strong>Household income</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than $30,000</td>
<td>45</td>
<td>8</td>
<td>37</td>
</tr>
<tr>
<td>$30,000–$49,000</td>
<td>75</td>
<td>15</td>
<td>60</td>
</tr>
<tr>
<td>$50,000–$75,000</td>
<td>90</td>
<td>23</td>
<td>67</td>
</tr>
<tr>
<td>More than $75,000</td>
<td>93</td>
<td>32</td>
<td>61</td>
</tr>
</tbody>
</table>

Table 1  Growth in adult U.S. Internet users, 1995–2006. Falling personal computer prices have contributed to rapid, if uneven, growth in Internet access in all major socioeconomic groups in the United States.

Source: Created by author, based on data from www.census.gov/compendia/statab/tables/08s1128.xls.
has access and who does not, particularly at home. In 1995 roughly one third of upper-income households (more than $75,000 annually) used the Internet; by 2006, this share had risen to 93%. Rapid growth rates also occurred among those of more modest means, although less than a majority (45%) of poor households (less than $30,000 annually) were users in 2006. Thus, as with race/ethnicity and educational level, absolute discrepancies persisted, but relative differences declined as Internet usage rates advanced most rapidly among those with hitherto the least access.

Schools remain perhaps the most important arena in which the digital divide is manifested and reproduced. In an age in which the acquisition of skills to participate in advanced producer services is key to upward social mobility, this issue assumes special importance. Inequalities in school funding are mirrored in the prevalence of the Internet in public classrooms: While 99% of schools offer children access to networked PCs in one way or another, these rates vary significantly in terms of quality of access. After home and school, public libraries are the third most common point of Internet access for children, especially for lower-income minorities; however, libraries have limited hours and often lack high-speed connections. Not surprisingly, the digital divide in schools has strongly racialized overtones: White students are much more likely than are minorities to use the Internet in the classroom or school library.

The latest frontier in the digital divide is unquestionably the arena of broadband delivery services. Broadband applications include digital television, business-to-business linkages, Internet gaming, telemedicine, videoconferencing, and Internet telephony. With large, graphics-intensive files at the heart of most Internet uses today (e.g., downloading forms, reading online newspapers), broadband has become increasingly imperative for Web browsing. In 2006, roughly 44% of the U.S. population used broadband technologies of one sort or another. This proportion is relatively low compared with most of the economically developed world; indeed, under the Bush administration, the U.S. slipped from 3rd to 13th internationally in terms of relative access to broadband services, and Americans pay 10 to 20 times as much per megabit for broadband as do their counterparts in Korea and Japan. Broadband accessibility in 2006 closely mirrored that of the Internet as a whole: It tends to be most prevalent among the young, males, whites, and the well educated and rises monotonically with household income. Such social differentials are accompanied by spatial ones. Broadband technologies have been slow to reach rural America: Whereas 86% or residents in cities with more than 100,000 residents have access to DSL, very few in towns with fewer than 10,000 people do so. Thus, there are strong reasons to believe that far from eliminating the digital divide, broadband reproduces it, gives it new form, and, in some cases, accentuates it.

Conclusion

Contrary to the hyperbole that continues to swarm around the Internet, multiplying even faster than do viruses and Web pages, cyberspace reflects all the inequalities and social divisions that permeate the nonvirtual world. Far from constituting some mythologized world of unfettered individualism, as some advocates portrayed it, cyberspace is in fact thoroughly shot through with relations of class, gender, ethnicity, and other social categories. Theorizations of the digital divide must of necessity take these dimensions into account to avoid the overly optimistic, technologically determinist, and often conservative perspectives that deny their ongoing existence and significance to understanding the Internet. When viewed in social terms, the interpenetration of the virtual and real worlds is mutually constitutive: Discrepancies in access to the Internet both mirror and constitute inequalities in the world outside cyberspace.

Barney Warf

See also Cyberspace; Telecommunications and Geography; Virtual Geographies

Further Readings

DIGITAL TERRAIN MODEL

Terrain models have always appealed to military personnel, planners, landscape architects, and civil engineers, as well as other experts in various Earth sciences. Originally, terrain models were physical models, made of rubber, plastic, clay, sand, and so on. Today, terrain models are represented in digital form, leading to the terminology digital terrain model (DTM). In this entry, first the basic concepts are introduced; next, the methodology for generating digital terrain models is described; and finally, the analysis and application of such a model are outlined.

Basic Concepts

A DTM is a digital representation of the terrain topography. The term was coined by Miller and Laflame of the Massachusetts Institute of Technology (MIT) in 1958. Their original definition was simply a statistical representation of the continuous surface of the ground by a large number of selected points with known X, Y, Z coordinates in an arbitrary coordinate field.

A number of terms related to DTM have been in use. These include digital elevation model (DEM), digital height model (DHM), digital ground model (DGM), and digital terrain elevation model (DTEM). These terms originated from different countries. DEM was widely used in the United States; DHM came from Germany; DGM was used in the United Kingdom; and DTEM was introduced and used by the U.S. Geological Survey (USGS). They have slightly different meanings. A DGM more or less has the meaning of “a digital model of a solid surface.” In contrast to the use of ground, the terms height and elevation emphasize the “measurement from a datum to the top” of an object. They do not necessarily refer to the altitude of the terrain surface, but in practice, this is the aspect that is emphasized in the use of these terms. Among these terms, DEM is nowadays the most widely accepted one for the digital representation of terrain topography, with terrain heights organized in a matrix form. The meaning of terrain is more complex and embracing. It may contain the concept of height (or elevation) but also attempts to include other geographical elements and natural features. Therefore the term DTM tends to have a wider meaning than does DEM and attempts to incorporate specific terrain features, such as rivers, ridge lines, and break lines, into the model.

Beyond DTM is a model called the digital surface model (DSM). It is a digital representation of the actual terrain surface. In other words, in a DSM, not only the terrain topography but also the spatial features (such as buildings, trees, and roads) on the terrain topography are represented.
If one removes all these spatial features from a DSM, one obtains a DTM. Indeed, the relationship between a DTM and a DSM may be compared with that between a contour line map and a topographic map.

Related to DTM, there are four basic issues: (1) how to acquire a set of data for the generation of DTM, (2) how to generate a DTM, (3) how to control the quality of the DTM, and (4) what can be done with DTM.

The process of construction of a DTM surface is called digital terrain modeling.

**Data Acquisition for DTM Generation**

Data acquisition is the collection of data to answer three basic questions:

1. From what sources (data sources) are data to be acquired?
2. What kinds of points are to be measured?
3. What techniques should be used for measurement?

Data sources are the materials from which data can be acquired. For the generation of a DTM, the actual terrain surface is of course the direct source. Indirect sources include aerial and/or satellite (or space) images and existing contour maps. The images could have been taken using scanners, radar, or frame cameras. For the height information of terrain points to be obtained, there is a need of an overlap between two images (taken at slightly different positions). For aerial photographs taken by a camera, the standard overlap is 60%. Such images are called stereo images, and a three-dimensional (3D) model can be reconstructed only in the overlapping area. For radar images, both the amplitude image and the phase angle are recorded.

To model a piece of terrain surface, first a set of data points needs to be acquired from the surface. In data acquisition, two stages are distinguished: sampling and measurement. Sampling refers to the selection of locations, while measurement determines the coordinates of the locations.

A number of sampling strategies have been in use. The most popular strategy is grid-based sampling. In this strategy, only those points that are located at the nodes of a designed grid are selected. Resultant height data can be easily organized into a matrix form. However, this doesn’t consider the importance of individual points, and, indeed, very important points (VIPs) may have been omitted.

In contrast to grid-based sampling is a strategy called selective sampling. In this strategy, VIPs are selected, such as peaks, pits and passes, points along ridges, courses, and break lines, and those points at which the change in slope is significant. A combination of grid-based sampling and selective sampling leads to a new strategy called composite sampling. Some other sampling strategies are also in use, although they are not very popular, such as progressive sampling (a type of grid-based sampling, with the grid made progressively finer) and contour-based sampling (i.e., selection of points along contour lines). After sampling, three attributes may be used to measure the quality of the data, that is, accuracy, distribution, and density.

Field surveying instruments such as total station and GPS (global positioning systems) receivers may be employed for measurement of the 3D coordinates of any points directly from the terrain surface. Field surveying is feasible only for a small area, although high-quality (or high-accuracy) data can be obtained. Photogrammetry is a technique used for measurement/computation of 3D coordinates from stereo images (or photographs). Traditionally, a pair of stereo images are used to reconstruct an analog model, and the measurement is carried out on the reconstructed analog model. The reconstruction is a process to restore the six orientation elements (three for position and three for angular rotations) of each image. In the digital era, the measurement has been replaced by computation. If the positions (x, y coordinates) of a point on both images (of the stereo pair) and the orientation elements of each image are known, then the 3D coordinates of the point on the ground can be computed using a mathematical model. Nowadays, for a given point on one image, the search for a correspondence point (conjugate point) on the other image has also been automated by a technique called image matching. With matching techniques, the acquisition of data from digital images for DTM generation has been fully automated. In radar (more precisely synthetic aperture radar—SAR)
images, when only the amplitude images are used, the principle is similar to that of photogrammetry, although different mathematical models are used for the computation due to the different imaging geometry. The technique is called radargrammetry. If, on the other hand, the phase angle information of the SAR images is used, then interferograms can be generated, and the heights are computed from the information contained in the interferograms. This technique is called SAR interferometry or interferometric SAR (InSAR). Cartographic digitization from existing contour line maps using digitizers or scanners is employed. If scanning is used, then a raster-to-vector conversion is required to convert the scanned contours into vector-based contour lines.

Modeling Approaches for DTM Surface Reconstruction

The process of reconstruction of a DTM surface is called digital terrain modeling. It is also a process of mathematical modeling, that is, fitting a mathematical surface onto the set of acquired data. Triangle-based and grid-based modeling are the two approaches commonly used. In the former, the set of data points is first triangulated into a triangular irregular network (TIN), and then a surface is fitted onto the TIN. If the surface determined by each triangle is used to represent only the area covered by the triangle, then the whole DTM surface can be formed by a linked series of contiguous triangles. It is also possible to fit a polynomial of order 2 or higher into the TIN to reconstruct a smooth surface. In grid-based modeling, if a surface is reconstructed from four grid nodes only, then the result is a bilinear surface; however, it is normal practice to fit a polynomial of order 2 or higher onto a large set of nodes to form a smooth surface over a large area.

In the case of triangle-based modeling, the DTM surface can be reconstructed directly from acquired data. However, in the case of grid-based modeling, if the acquired data are not in a grid form, then an interpolation is applied to the acquired data to form a regular grid and then the surface is reconstructed from the grid data. Such an interpolation process is often referred to as random-to-grid interpolation.

Interpretation of DTM

To interpret a DTM means to understand the terrain characteristics through the extraction/computation of the parameters. DTM interpretation is also called DTM-based terrain analysis.

The term terrain analysis means different things to people with different backgrounds because they emphasize different aspects. In some literatures, the emphasis in digital terrain analysis is on interpolation methods for terrain surface modeling; and in some other literatures, the emphasis is on visualization of DTMs. However, to most people, it means the derivation of attributes from terrain surfaces.

Some parameters related to the geometry, morphometry, hydrology, and visibility of a terrain surface can be computed from a DTM. The geometric parameters include the surface area, projection area, volume, and so on. Morphometric parameters are those that can be derived directly from the DTM using some local operations, such as slope and aspect, rate of change in slope and aspect, plan and profile curvatures, roughness, and so on. Flow lines, catchments, and hydrological networks can be extracted from a DTM. Two fundamental parameters in visibility analysis can also be computed from a DTM, that is, point-to-point visibility (intervisibility of line of sight) and point-to-area visibility (viewshed).

These parameters are also called primary topographic attributes. A number of other secondary topographic attributes can be computed from them, such as wetness indices, stream-power indices, radiation indices, and temperature indices.

Quality and Accuracy of DTM

Quality, efficiency, and economy are three critical parameters in digital terrain modeling. The question of how to measure the quality of a DTM is not easy to answer. In general, two types of quality measures are suggested, that is, the producer’s measures and the user’s measures. DTM users may demand that the quality of terrain parameters computed from a DTM should be used as a measure of the DTM quality. This insistence is reasonable. However, this approach requires a larger number of measurements to be used, and this is not feasible in practice. The producers of
DTMs are national mapping agencies, and they make use of the root mean square (RMS) errors at check points as a quality measure, following the tradition of topographic maps. Such a quality measure is referred to as DTM accuracy. If there are no height errors in a DTM, then all terrain parameters computed from the DTM are also free of error.

It has been found that the RMS error of a DTM increases with greater intervals between acquired data points. The trend is quite linear if a composite sampling strategy is employed for data acquisition. However, the trend will be of the second order if only grid-based sampling is employed for data acquisition.

Another issue related to the quality of DTMs is quality control. This refers to the precautions or actions taken to ensure the quality of the final DTM. Quality control could be applied online and/or offline. Online quality control means making quality inspections during data acquisition and/or surface modeling to avoid gross and/or systematic errors. Offline quality control means detecting gross errors and rectifying systematic errors after data acquisition and surface modeling. Various gross error detection techniques are available.

Applications of DTMs

DTMs have recently become an important part of national geospatial data infrastructure. DTMs has found widespread uses in all geosciences and engineering, such as

- planning and design of civil, road, and mine engineering;
- 3D animation for military purposes, landscape designing, and urban planning;
- analysis of catchments and hydraulic simulations;
- analysis of the visibility between objects on the terrain surface;
- terrain analysis and volume computation;
- geomorphological and soil erosion analysis;
- remote sensing image interpretation and processing; and
- various types of geographical analysis.

Zhilin Li

See also Aerial Imagery: Interpretation; Cadastral Systems; GIS in Land Use Management; Landscape Architecture; Landscape Quality Assessment; Land Use Analysis; Map Animation; Map Visualization; Scale in GIS; Suitability Analysis; Terrain Analysis; Three-Dimensional Data Models

Further Readings


Digitizing or digitization is a process that represents an object, image, or document by a discrete set of points, lines, polygons, or raster images. The result is called a digital representation or a digital image of the object represented. In geographic information systems (GIS), digitizing is often the process of converting features on paper or an analog map into digital format. To digitize a map, a GIS practitioner will often use a digitizing tablet (also known as a digitizer) connected to a computer to trace over the features of interest or a scanning device that creates a digital image that is later manipulated in the GIS software.

When using a digitizing tablet, the $x$, $y$ coordinates of the features are automatically recorded and stored as spatial data after the paper map is registered in a coordinate system that the user designates. Digitizing with a digitizing tablet
offers another way, besides screen digitizing “freehand” (also called “heads-up” digitizing), to create and edit spatial data. Users can convert features from almost any paper map into digital features. Typical devices used for digitizing are a digitizing board or a tablet and a cursor. The digitizing board consists of tiny wires that run horizontally and vertically, forming a grid. The digitizing cursor has an optical viewer with crosshairs, often called a puck, which allows the user to visually locate a point on the map. The most common cursor has at least 16 buttons and contains a user-definable keyboard. Digitizing converts spatial features on a map into a digital format. Before actually starting to digitize, a map is mounted on a digitizing board and then registered on a coordinate system, and a digitizer cursor is used to trace each map feature. When a point is identified on the map, the user presses a button on the cursor and the computer records the current x, y coordinates of that position in the designated digitizer units. These become the x, y coordinates of the point feature or one of the points comprising a line or polygon. The location of the digitizer cursor’s crosshairs on the board and the map is simultaneously displayed on the graphic display screen. In addition to using a digitizing tablet, large-format scanning devices are also used to create digital representations of maps, which are then “vectorized” and edited within a GIS application.

Users often use a digitizer in conjunction with editing tools in GIS applications such as ArcMap or MapInfo to create new features or edit existing features on a digital map. The user can digitize features into a new GIS data set and then add additional data to an existing map once it is in a digital format. Cadastral mapping is a common task involving a digitizing tablet that traces property parcels drawn on photography printed on Mylar or paper. This process allows the user to create a digital representation of property parcels that can be updated and/or edited more easily. Furthermore, the digital representations can be used to perform spatial analyses that would be more cumbersome using paper or Mylar maps.

Shawn Lewers

See also GIScience; Vectorization

Further Readings


Disability, Geography of

The geography of disability is an important and emerging area of human geography scholarship. Research in this area of the discipline has increased over the past two decades, to the point that it is now a distinctive subdiscipline referred to often as the geographies of disability. The nature of research in this subdiscipline has greatly evolved over time due to changes in research interests, the infusion of ideas from other disciplines such as disability studies, the use of a greater range of research methods, and the development of stronger connections with other geographic subdisciplines such as health geography. Much research in the geographies of disability is united by a shared interest in understanding sociospatial experiences of disability and processes of disablement. This entry gives an overview of models of disability and examines research on the geography of disability, including emerging work on mind and body differences as well as issues of research design and dissemination of concern to the subdiscipline.

Central to all disability research, including that undertaken by geographers, is understanding what disability is; there is, however, no single, definitive way of doing this. Rather, there are multiple models to assist with characterizing disability. The key difference between these models centers on where the “cause” of disability or disablement is located. Biomedical models of disability identify mind and body differences (often expressed as physical and mental impairments) as the source of disability. In such a model, disability is based on individual deficiency (i.e., an inability of one’s body and/or mind to work in expected and “normal” ways). Social models locate the causes of disability as external to the body and mind, often attributing disabling experiences to social structures, systems, and institutions that constrain people’s abilities to live in
desired ways. Numerous other models of disability exist, including the biopsychosocial and social justice models. There are debates within human geography and beyond as to which model is most appropriate to use in research. From this brief overview, it can be understood that what constitutes disability varies quite significantly, based on the model that is employed. The particular model adopted by a geographer to guide his or her work in the geographies of disability subdiscipline will very much inform the research question(s) ultimately posed. The adoption of different models of disability by geographers within the subdiscipline has, in part, contributed to the breadth of topics investigated.

Human geographers in this subdiscipline have explored a range of research topics that have changed in focus over time. An early area of inquiry, for example, was related to physical access to the built environment for those with a range of bodily impairments. More recently, theoretical developments in human geography and beyond, the evolution of models of disability, and the emergence of the discipline of disability studies have all contributed to a broadening of areas of inquiry. Recent years have seen the emergence of research in the subdiscipline regarding the enabling and disabling nature of technologies, sociospatial barriers to desired participation in society and space, the embodiment of disability and impairment, disability activism and engagement, and work and leisure spaces occupied by disabled people.

The increasing breadth of inquiry in the geographies of disability has also resulted in a greater range of mind and body differences being explored in research, particularly in the past decade. Chronic illnesses, for example, are diseases, disorders, and syndromes that are not curable and have symptoms, and resulting impairments, that typically fluctuate and may be invisible. People living with such illnesses can be constructed as disabled and/or experiencing disablement, depending on the model of disability employed. Chronic illness is an area gaining increasing attention in this subdiscipline, and such research is often situated at the intersection of geographic research on both disability and health. Other mind and body differences rising in interest in the subdiscipline include mental illness and learning disabilities. Emerging areas of study in human geography more broadly are also developing foci on mind and body differences that have implications for how we understand disability, including within the geographies of disability subdiscipline, such as research related to environmental illness, obesity and fatness, and cancers in that these can all have disabling outcomes for individuals.
The power dynamics inherent in conducting research with marginalized populations, and a desire to conduct research with and for disabled people and not on them, have brought forth attentiveness to issues of research design and dissemination within this subdiscipline. The ethics and politics of knowledge production, accessibility and usefulness of research findings to disabled people and their struggles for inclusion, and the achievement of desired social and political outcomes resulting from the research process and uptake of findings, for example, are issues that have gained considerable attention within recent years. Not surprisingly, debates regarding such facets of the research process are not isolated in the subdiscipline but rather borrow from and contribute to those of disability scholars across disciplines. Another issue increasingly gaining the attention of human geographers involved in disability research is that of the potential for developing participatory processes in research. Doing so serves as a response to the desire of some community members and advocates to serve as partners in research and a response to academics’ increasing interests in developing useful working partnerships with participants/co-researchers as a way of partially undoing power hierarchies.

Valorie A. Crooks

See also Accessibility; Blindness and Geography; Body, Geography of; Difference, Geographies of; Health and Health Care, Geography of

Types of Disaster and Prediction Methods

Disaster Types

Disasters are classified as natural or technological. Natural hazards include floods, earthquakes, hurricanes, storm surges, tornadoes, wildfires, landslides, tsunamis, volcanic eruptions, severe winter storms, droughts, extreme heat, coastal erosion, thunderstorms, hailstorms, snow avalanches, land subsidence, expansive soils, and dam failures. Technological hazards include fires, hazardous material accidents, nuclear plant accidents, terrorism, and biological or chemical weapons. Not all disasters are equally predictable. Hurricanes can be predicted several days in advance, whereas tornadoes and earthquakes often take communities by surprise. Warnings therefore tend to be much more successful for hurricanes in comparison with tornadoes. Earthquakes are virtually unpredictable, and therefore, society cannot be warned with much success.

Prediction Technologies

Disaster warning systems and the communications technologies through which warnings reach
DISASTER PREDICTION AND WARNING

Communities remain limited without improvements in prediction technology. One goal of improving prediction technologies is to increase the lead time. Lead time is the amount of time between when a warning or alert is issued and when the hazard for which the alert was issued strikes the community. If, for instance, a warning was issued for a tornado at 10:00 a.m. and the tornado struck at 10:10 a.m., the warning generated only 10 minutes of lead time. Short lead times (the result of poor predictive technology) give communities little time to prepare, which, in turn, leads to increases in the number of fatalities and injuries. Tornadoes are one disaster for which prediction technologies are rapidly evolving.

The Warning Decision Support System (WDSS) is an example of a new technology that is also communication friendly. WDSS applies automated algorithms through the WSR88-D Doppler weather radar system to automatically detect and predict severe weather, revealing weather conditions using the color display system that is now widely known to the public and used by weather forecasters in the communication process. The Automated Weather Interactive Processing System (AWIPS) serves to integrate the data that come from Doppler, ASOS (Automated Surface Observing Systems), GOES-8 and GOES-10 (Geostationary Operational Environmental Satellites), wind profilers, and other observing systems.

Recently three-dimensional (3D) numerical modeling studies of tornadoes have been carried out. These models show a tornado’s structure and wind field. Beyond providing information on the formation of a tornado, the Doppler system has increased the lead time for tornado warnings. These systems can be set so that an automatic alert goes off in the forecast office whenever a storm, within that station’s radar umbrella, exceeds a given value of reflectivity. The challenge lies in adjusting the 120 geographically dispersed Doppler systems to local climatologies and developing new algorithms for tornado detection and warnings. In addition, prediction can be improved with 3D numerical prediction improvements.

The Warning Communication Process

Successful disaster predictions, however, do not always lead to effective warnings. To be truly
effective, hazard alerts must meet the needs of those participating in the communication process. Warnings should take into account goals, participants, processes, and technologies available to local communities.

**Notification Types and Goals**

The goal of disaster communication is to inform the public of the danger and the means by which threats can be avoided. The details of warning communication goals vary depending on the types and degrees of hazard in a disaster. For instance, with tornadoes, two levels of notification are employed: a “watch”—meaning that conditions for a tornado are present—and a “warning”—meaning that a tornado is about to occur, is presently occurring, or has occurred. Depending on the type of notification, different goals and processes will be used in communication.

**Participants in the Communication Process**

For a disaster warning communication to be efficient and effective, all participants in the process must be considered. The general public, federal agencies (such as the National Weather Service, NWS), meteorologists, forecasters, news media, public officials, local emergency personnel, and volunteers are all involved in the process of disseminating and receiving pertinent disaster notifications. Coordination strategies must take into account the different structures and cultures of these organizations.

**Steps of the Process and Technologies Employed**

Due to the complexity of disaster situations, the steps of the communication process are not always linear and systematic. Forecasting and prediction agencies, both public and private, work in partnership to share prediction information and assessments. Federal, state, and local agencies will often converge and communicate among themselves prior to issuing warnings to decrease the likelihood of providing contradictory information to the public. In the case of the NWS, all levels of government must coordinate their efforts in warning the public. Local, county, and state emergency managers cooperate with the NWS in the prediction process. The media then disseminate the information using television, radio, and Web sites. The NWS may activate tone-activated NOAA (National Oceanic and Atmospheric Administration) radios in the United States, whereas local emergency managers might sound sirens or activate “Reverse 911” to warn the public. Reverse 911 is a telephone communications alert system that uses a patented combination of database and geographic information systems mapping technologies to deliver outbound notifications.

**Interpretations of Warnings**

Even if a disaster warning is issued timely and effectively, the public still faces the problems of interpreting the information that the warning intends to convey. Interpretation takes place on many levels; however, there are two important trends in how the public interprets warnings. The first of these is the influence that social stratification has on a warning’s interpretation. Second, local and popular myths about the dangers of disasters may lead some people to disbelieve warnings. Both these factors influence how and why people respond to disasters.

**Social Stratification and the Receipt and Interpretation of Warnings**

Forecasts inform responders and the public as to where, when, and with what level of severity a disaster is likely to occur. Eliciting a response requires forecasters to know how communities perceive and react to weather forecasting and warnings. Warnings only save lives and prevent injuries if they are understood, meet individual needs, and are accurate, reliable, and timely. In reality, these ideal conditions are difficult to meet: rates of poverty, education, gender roles, race, ethnicity, culture, and previous experience with weather among communities limit the effectiveness of public forecast information. People who have the most access to forecasts, for example, are usually the educated and affluent and tend to be members of majority groups. Assuming that the forecast
has been heard, the forecast still needs to be understood, believed, confirmed, and responded to, all of which depend on public demographics, social, cultural, economic, and psychological patterns. Minorities, for example, are less likely to adopt an emergency plan because they receive the warnings less frequently.

Weather Web sites are the primary form of weather information. However, not all groups have equal access to computers—the elderly, for instance. Another emerging form of disaster communication is Reverse 911, as noted earlier. This, too, is a stratified warning system as many young adults do not have landlines and rely on cell phones.

Jamie Mitchem studied the effects that a tornado can have on an urbanized area. The tornado hit Marion County, Indiana (which includes the state capital, Indianapolis), which had a population of 860,454 in 2000. Eleven percent of the population was above the age of 65. Of all the characteristics of the people within the county, the most striking was that 157,908 had disabilities. People with disabilities are most susceptible to disasters. When the tornado hit Marion County, 47% of respondents never received a warning. Of the respondents who were aware of the tornado, 81% received warnings. Of that 81%, only 31% felt that they were in real danger. Of the people who received warnings, 27% ignored the warnings, and 49% did not seek shelter. Of all the respondents, 70% knew what a tornado watch meant; 21% were unaware of the definition of a tornado warning; and 22% had never performed a tornado drill and lacked a family emergency plan. The most effective forms of warning were tornado sirens and local television networks.

Myths and exaggerations concerning disasters and how people typically respond to them are perpetuated by the media and by Hollywood films. The media, for its part, tends to overstate the number of incidents of looting, postabandonment by officials, price gouging, and public panic. In actuality, some researchers suggest that crime may actually decrease after a disaster has struck. Price gouging is rare in disaster situations, and most behavior during disasters is altruistic in nature. The perpetuation of these myths and exaggerations can be detrimental to the communication and response process. It is important that they be debunked for officials, volunteers, and the media to improve public response rates.

Hollywood, rather than exaggerating actual events through reporting, perpetuates myths by repeating stereotypical disaster responses. Dante’s Peak, Volcano, and Asteroid serve as key examples within Hollywood’s disaster genre. In these films, lawlessness and violence are presented as how humans normally respond to disasters. These films also promote erratic and self-centered behavior as typical and expected. Research in sociology and other social sciences regard this typification as false.

Myths and exaggerations also distort the types of behaviors that people in actual disaster situations should expect from one another and could decrease the likelihood that the public will follow response tactics and take appropriate action at appropriate times. For example, some people may refuse to evacuate for fear of looters. It is vital that emergency managers and volunteers understand how people behave in disaster situations.

William Donner, David Cohen, and Marci Cottingham

Influence of Myths and Exaggerations on Interpretation

Myths and exaggerations concerning disasters and how people typically respond to them are perpetuated by the media and by Hollywood films. The media, for its part, tends to overstate the number of incidents of looting, postabandonment by officials, price gouging, and public panic. In actuality, some researchers suggest that crime may actually decrease after a disaster has struck. Price gouging is rare in disaster situations, and most behavior during disasters is altruistic in nature. The perpetuation of these myths and exaggerations can be detrimental to the communication and response process. It is important that they be debunked for officials, volunteers, and the media to improve public response rates.

Hollywood, rather than exaggerating actual events through reporting, perpetuates myths by repeating stereotypical disaster responses. Dante’s Peak, Volcano, and Asteroid serve as key examples within Hollywood’s disaster genre. In these films, lawlessness and violence are presented as how humans normally respond to disasters. These films also promote erratic and self-centered behavior as typical and expected. Research in sociology and other social sciences regard this typification as false.

Myths and exaggerations also distort the types of behaviors that people in actual disaster situations should expect from one another and could decrease the likelihood that the public will follow response tactics and take appropriate action at appropriate times. For example, some people may refuse to evacuate for fear of looters. It is vital that emergency managers and volunteers understand how people behave in disaster situations.

William Donner, David Cohen, and Marci Cottingham

See also Coastal Hazards; Differential Vulnerabilities to Hazards; Disaster Preparedness; Earthquakes; Floods; GIS in Disaster Response; Hurricane Katrina; Natural Hazards and Risk Analysis; Three-Dimensional Data Models; Tornadoes; Tsunami of 2004, Indian Ocean

Further Readings

Disaster preparedness consists of those actions taken by individuals, organizations, or communities to ready themselves to respond to and recover from emergency situations. As the word *preparedness* may indicate, these are actions taken prior to an impact, when disaster preparedness is in reality getting ready for a *hazard*, which is a potential threat to society. Disasters are events that cause a significant disruption to society or the environment. Risk, vulnerability, and resilience are related concepts in that they refer to a community’s susceptibility to or capacity to withstand or bounce back from future hazardous events. Disaster preparedness is important because the better prepared an individual, organization, or community is for such an event, the smaller the disruption ought to be. Disaster preparedness is one phase of emergency management. It has also been a topic of research for geographers interested in studying the human response to hazards, a field pioneered by Gilbert White, the father of floodplain management and founder of the Natural Hazards Research and Applications Information Center. This entry briefly discusses disaster preparedness from both emergency management and hazards geography perspectives. The impact of research on human response to hazards has the ability to inform public policy, and the emergency management system provides the context for community-, state-, and national-level decision making and is an important subject for future research.

**Emergency Management Perspective**

In the United States, the Federal Emergency Management Agency (FEMA) is responsible for government action in preventing, preparing for, responding to, and recovering from disasters. FEMA was formed in 1979 as a merger of several agencies with a diverse range of responsibilities, including the Federal Disaster Assistance Administration and the Defense Civil Preparedness Agency. Much of the history of emergency management was characterized by response to disasters, rather than preparedness activities. Early attempts to prevent damage and loss of life from natural disasters included the Flood Control Acts of the 1930s through 1950s, which established and defined the government’s role in flood control and authorized flood preparedness and emergency operations. The National Flood Insurance Program, created in 1968, was also an effort to lessen the impact of flood hazards and prevent losses from floods. The Civil Defense program was one of the most important influences on the need to prepare for emergencies. Part of its efforts to prepare for a Cold War attack included education campaigns and the development of evacuation plans. Preparedness and mitigation activities for other hazards became more important in the 1990s as the Cold War ended. The Disaster Mitigation Act of 2000 required that states maintain mitigation plans as a condition for receiving disaster assistance. In 2003, 2 years after the terrorist attacks of September 11, FEMA became part of the Department of Homeland Security, and a renewed prominence of preventing, preparing for, and responding to willful disasters was added to dealing with natural and technological hazards.

Governmental organizations similar to FEMA exist in other countries, such as Public Safety Canada or National Disaster Management of India. Also, the United Nations’ International Strategy for Disaster Reduction aims to reduce disaster-related losses through increasing awareness of the importance of disaster reduction as a component of sustainable development. The World Health Organization also maintains a disaster preparedness and response program to assist countries in preparing for disasters.

Disaster preparedness is one of the four phases of an emergency management model,
along with mitigation, response, and recovery (Figure 1). Mitigation involves long-term practices that prevent or lessen the impact of disasters and includes actions such as improvements in forecasting and warning technologies, land use regulations, and building codes. The National Flood Insurance Program is considered mitigation. Preparedness activities include actions taken to prepare for an expected event, including gathering supplies for a disaster kit, making an evacuation plan on the part of a household and conducting drills, and acquiring resources to aid in disaster recovery efforts on the part of communities. Response and recovery occur during and after the disaster and include search and rescue, evacuation or sheltering in place (response), and reconstruction and rebuilding of infrastructure (recovery). The four phases are interrelated, non-mutually exclusive, and cyclic. The post-impact desire to prevent future losses can encourage mitigation efforts that might otherwise have a difficult time gaining acceptance. Recent events can also encourage preparedness for as well as response to subsequent events.

Preparedness actions for individuals and households generally involve being ready either to evacuate or to shelter in place, depending on the location and the event. Items that are often recommended to have at hand in the event of a disaster include several days’ supply of food and water, a battery-powered radio and flashlight, extra batteries, a first-aid kit, sanitation items, a can opener, and cash. Families with special needs or animals have other requirements, including medicine, diapers, and pet supplies such as pet food and a means to transport an animal. In addition to a preparedness kit, preparedness actions are also encouraged, such as identifying an out-of-town contact that family members can communicate with and planning ahead of time where one would go in the event an evacuation was required.

Community-level preparedness is concerned with the immediate response to the event as well as the preparedness of the community for the recovery process. Research has identified four basic functions that emergency management must perform. Preparedness begins by assessing the demands a disaster will place on a community and its ability to respond in those four basic ways, which include (1) an assessment of the possible impacts of the event, (2) last-minute mitigation efforts to protect property such as sandbagging, (3) population protection through actions such as evacuation orders, and (4) management of incident personnel and resources. These four functions identified by research are similar to the initial response objectives specified by the Department of Homeland Security’s National Incident Management System, which are saving lives, incident stabilization, and property preservation. Planning for the recovery process is important because more planning can mean a more efficient allocation of resources and better odds of full recovery.
Historically, much of hazards and disaster research has been on natural (e.g., hurricanes, earthquakes, floods) hazards and their impacts; later, there has been research on technological (e.g., toxic waste, nuclear power accident) hazards. Willful hazards (e.g., terrorist attacks) and disasters have recently gained greater attention from researchers. Much of the existing literature has been devoted to preparedness and mitigation for individual hazards, although there are several publications that attempt to address decision making across hazards. Preparedness includes those actions taken by individuals and households as well as those actions taken by organizations or communities. Most research tends to focus on a single level, especially individual or household-level decision making. Recent work has examined decision making in organizations or private businesses. Research generally agrees that hazards and disaster preparedness are not typically given high priority by individuals, organizations, businesses, and even communities until some event threatens or until action is mandated.

Research on individual or household-level disaster preparedness has focused on factors that make some individuals more likely to prepare for disasters, the types of adjustments they are likely to make, and the time or money they are likely to spend on preparations. Decisions are generally made by households but are situated within larger social groupings. Community norms can influence the range of preparedness activities seen as available options. Communities holding attitudes such as the belief that disasters are “acts of God” and cannot be controlled may make preparedness decisions different from those of communities that acknowledge the potential of risk and believe that the event can be predicted. Other factors associated with whether or not one takes any preparedness actions can be classified broadly as situational or cognitive factors. Cognitive factors can be either attitudinal or psychological and include individuals’ personality characteristics, perceptions about the hazard, perceptions about the preparedness action, or beliefs about personal vulnerability. Situational factors include the physical environment (characteristics of the hazard) and the socioeconomic environment (e.g., education and income).

An individual’s socioeconomic status can be a limiting factor to taking actions to prepare for a disaster. Poverty limits one’s ability to make some adjustments that can reduce future impacts from disasters. Income, education, and home ownership have also been related to making preparations for specific hazards, with higher levels of these being associated with higher levels of preparedness or a greater number of preparedness actions. Psychological and attitudinal factors commonly associated with preparedness include self-efficacy, which is the belief that one has the ability to carry out successfully the protective action, perceived personal responsibility, and perceived threat of the hazard, which can be influenced by many factors itself. Prior experience with the hazard has also been associated with preparedness and mitigation decisions.

Less research has been conducted on community- or business-level preparedness than on household-level preparedness. Communities with repeated and recent exposure to hazards tend to be more organized. Legal incentives and mandates have been responsible for some preparedness efforts. The Mitigation Act of 2000 noted earlier is one example of mandating preparedness plans as a precondition to receiving federal disaster assistance. Another is a Louisiana state law requiring businesses that care for companion animals to file emergency plans. Size, access to resources, and awareness of the hazard are factors that have been associated with preparedness by communities or businesses.

Several strategies to study preparedness activities have been used by different disciplines throughout decades of research. Utility models, in which individuals make decisions to maximize all the possible outcomes of those decisions, provided a basis for early hazards research. Individuals cannot process all the information available to make a rational decision, however, and they bring other goals to the decision-making process. Decisions are made with bounded rationality and thus are based on subjective assessment of utility and result in less than optimal final decisions. Heuristics explain individuals’ decision-making behavior as the result of several common subjective assessments. Heuristics help place the decision making in the context of more familiar events through mental shortcuts and are important...
because people must make decisions based on new information under conditions of uncertainty. The availability heuristic, for instance, could lead individuals to believe that a particular event was rare if they could not recall any instances of such an event (and therefore they would not have to prepare for it). The general model of risk communication, which emerged through years of hazards research, stated that a message must be heard, confirmed, understood, believed, and personalized before action would be taken. Many of the ideas discussed above are complementary and acknowledge that individuals’ preparedness or mitigation actions are the result of internal and external factors and that a combination of factors either encourage or constrain actions. Most also operate under the assumption that an individual (or business or community) is not likely to prepare for an event unless he or she becomes aware of the event and believes that the event has consequences for him or her.

Kathleen Sherman-Morris

See also Differential Vulnerabilities to Hazards; Disaster Prediction and Warning: Floods; GIS in Disaster Response; Hurricane Katrina; Hurricanes, Risk and Hazard; Natural Hazards and Risk Analysis; Resilience; Tsunami of 2004, Indian Ocean; Vulnerability, Risks, and Hazards; White, Gilbert

Discourse is a particular claim about the relationship between words and the things they represent. Discourse theory began with Michel Foucault, the famous French philosopher, as a critique of the Enlightenment idea that science and scientific investigation—using observation, measurement, and hypothesis testing—act as a mirror to the world, thus providing us rational objective truth about the world. Discourse theory rejects the claim that science offers us objective knowledge of the world and claims instead that all objects—while certainly real in a material sense—are also discursive constructions. Discourse theory claims that an object cannot be understood as a mere object but can only be understood in context, that is, in relation to the conversations, or discourses, that surround it. Imagine a car: It is in one sense a vehicle to get from Point A to Point B. But it is at the same time an object embedded in myriad discourses—environmental, social, and political, among others that reach far beyond its mere meaning as a way to get to and from work.

Discourses must be understood as constellations of statements that are not free-floating or disembodied but that act in powerful ways to shape our conception of the world and the objects in it. Invariably, the play of discourse sanctions some forms of knowledge as normal or true while excluding others. Social institutions play an important role in the production and circulation of discourse—what we talk about, how we talk...
about it, and what context we choose, or are allowed by the rules of what is normal, to discuss an object in. This is what is meant by the proposition that all objects are at once objects in a material sense and social constructions, and that is why Foucault equated discourse with power, knowledge, and truth. These claims have profound implications for scientific explanation, education, social policy, and politics of social change.

Discourse theory in the main rests on two principles common to all science and to geography: selection and aggregation. It is, of course, true that any object possesses an infinite number of attributes. Consider, for example, a human body, which presents itself to us as a single, tangible, bounded, objective unit of reality. But a human body has innumerable aspects such as physiology, personality, sexuality, medical history, and identity characteristics within each of which is another endless list of metric and nonmetric attributes. For a doctor, a body is a site of healing and surgery. For a biologist, a human body is a site of cellular processes. To a soldier, a human body is a target. To a census taker, it is a statistic; to a lover, a site of pleasure, and so on ad infinitum.

So what exactly is a scientific definition of a human body? Every description of a body is by necessity a partial description of attributes that have been selected from an infinite list. In other words, the body is discursively constructed for a purpose, and there are as many such constructions as there are purposes. Consider the proposition “The man is black.” A person’s skin color does not describe something fundamental or essential about him or her. The statement represents a discursive construction of a particular human body relative to the powerful body discourses that are prevalent in society. The constructivist claims I have made about the body of a single person are true for all objects of scientific inquiry ranging from a grain of sand to the near-infinite realms of the known universe. From an infinite number of attributes, we select certain elements; this is a natural and essential process, without the ability to select certain attributes and ignore others, conversation of any sort would be rendered impossible. The claim of discourse theory is simply that what we select, and what we ignore, are shaped by social forces and certainly cast doubt on the rational scientific claim to objective knowledge free from discourse.

The second constructivist principle is discursive aggregation. Every word is not only a partial description of the objects it represents but also a discursive aggregation of disparate things. Words group things together to enable conversations. Consider the word *capitalism*. That word has no objective, discrete, bounded, scientific entity that lies at the core of the word. The term *capitalism* groups a number of disparate things, including ownership over the means of production, flows of money, wages, advertising, pollution, personal consumption, health, nutrition, lifestyles, cultural values, and innumerable other elements. It represents a way of talking about the world at particular levels of discursive aggregation. The term does not represent a pregiven segment of reality but is simply a useful way of talking about some aspects of the social world. Discourse theory does not demand that we abandon such aggregate concepts but only that we be aware of aggregation and its implications for scientific study and conversation.

All objects, and the names we use to identify them, are by necessity aggregations of other objects. Consider the terms *molecule, soil, tree, forest, capitalism, deciduous forest, global warming, and terrorism*. We obtain the concept of a “deciduous forest” by selecting the characteristic of “annual leaf fall” out of the infinite number of attributes possessed by trees. The concept does not correspond to a bounded entity existing in the natural world; it is simply a way of representing a set of trees at a particular level of discursive aggregation.

The fundamental claim of discourse theory is that the relationship between words and the things they represent is always mediated by the two principles of selection and aggregation and that thus the “objective” world of science is in fact a discursive construction. There is no known methodology to escape discursive construction.

Application of Discourse Theory to Geography

How does the notion of discursive construction apply to geography—to its objects of study and
its concepts of space and scale? Unlike systematic sciences such as economics and geology, geography is not defined by a unique set of objects. The nearest formulation to the concept of a geographic object is “anything that has a specified location.” Such objects are shown on maps as points, lines, and areas, but the interest of geographers in them arises from their ability to represent substantive features such as cities, roads, and regions. Following the two principles invoked earlier, it is clear that all geographic objects are discursively constructed because the actual description of cities, roads, and regions cannot escape the rules of attribute selection and level of aggregation. Depending on the context, a city, for example, is more than just its population size; it is also defined by a number of other attributes such as air pollution, safety, lifestyle, and so on. It is also a discursive aggregation—of people, neighborhoods, streets, and traffic.

The rules of discursive construction apply with equal force to the concept of geographic space. Traditionally, geographers have viewed space as an a priori given (alongside time) that exists prior to geographic investigation, where the subject was in fact defined as the study of the organization of objects in space. The notion of space as a pregiven dimension is also evident in the literature on time-geography and space-time cubes. But why are geographers concerned with the concept of space? Independent of objects there is nothing intrinsically interesting about space itself. This provides us with a succinct definition of space: “Space is that in which objects are located.” Space is not an a priori dimension; it comes into being only through objects, their attributes, and their spatial relationships. But if objects are discursively constructed and space comes into being only through the location of objects, then elementary logic leads us to the proposition that space is also a discursive construction.

The geographic concept of scale provides a powerful insight into the role of discourse in geography. At its most elementary level, a scale is a ratio between a linear unit on the map and the number of ground units it represents. The appearance of detail of a city drawn at a scale of 1:5,000 is very different from what is depicted of the city at 1:500,000. The transition involves the selection of features and attributes to represent. Therefore, every map, drawn at whatever scale, is a discursive construction of the world. Geographers also use the concept of scale to speak about places as they range over the levels of households, neighborhoods, communities, counties, regions, and continents. Those conversations take place at particular levels of discursive aggregation. In fact maps of landforms, climate, or vegetation classifications are nothing but fact groupings based on the principles of selection and aggregation at different levels of generalization.

The written records of geographic knowledge dating back to the Greeks and Arabs show that they routinely employed concepts of space, scale, and classification. So the role of discourse was always present in geography and is not something that simply surfaced in “postmodern times.” It is manifestly clear that no geographic knowledge can exist outside discourse, but the more important point is to explore the significance of that relation.

The importance of discourse to geography can be illustrated by considering the common practice of drawing “poverty maps.” A household is considered poor when its income falls below a specified income threshold. Areas with high concentrations of poverty households are deemed “poverty areas.” Such areas are then compared with other areas of high income, permitting geographers to extract a series of explanatory indices such as unemployment, race, gender, type of household, education level, and behavioral profile. The list of such indices determines how social policy will be formulated and, consequently, who will be authorized to address the poverty problem. The extraction of explanatory indices based on the spatial distribution of household income is not an objective act of scientific investigation that simply mirrors a given underlying social reality but is, rather, a profound discursive construction of the material conditions of people’s lives. Poverty is not an inherent characteristic of certain countries, states, or even neighborhoods. Furthermore, the common practice of “mapping poverty” may tell us where poor people live but offers us no location for where the causes of poverty reside. As reasonable as the geography of poverty may at first seem, as a discourse it is linked as a causative agent of the very problem it is designed to address. By comparing poor areas with those
that are not, poverty maps set up an unhelpful binary along the following lines: poor versus non-poor, problem versus nonproblem, undeveloped versus developed, inner city versus suburbs, areas without resources versus those that have, poor as the other versus nonpoor as self. Such binaries privilege nonpoor areas and their residents, houses, landscapes, and lifestyles while ignoring the role of wealthy areas in creating poverty. By presenting the income levels and consumption habits of nonpoor areas, not just as the location of the nonproblem but in fact as the standard to emulate and aspire to, the geography of poverty has become a prescription for the creation of permanent scarcity and inequality.

Presenting the geography of poverty as discourse rather than as objective science not only allows us to contest that particular representation but also allows us to create alternative discursive constructions that may prove more useful. Alternative constructions will select different features to highlight, present them at different levels of discursive aggregations, authorize different social policies, and thus empower different agents to act. If we move away from the concept of income—on which poverty maps are based—and ask instead a series of alternative questions, such as why poor people find it difficult to obtain good nutrition, live in healthy bodies, or find affordable transport to work, the new discourse we generate leads us to very different solutions, social policies, and social change agents empowered to act. This is what Foucault meant by employing the neologism “power/knowledge”—that power is exercised through knowledge; contesting established discourses is indeed a form of political practice because it enables the exercise of power by different agents.

Lakshman Yapa

See also Art and Geography; Cultural Geography; Cultural Turn; Epistemology; Ethnocentrism; Eurocentrism; Film and Geography; Geographical Imagination; Humanistic Geography; Orientalism; Phenomenology; Photography, Geography and; Positionality; Postmodernism; Poststructuralism; Representations of Space; Spaces of Representation/Representational Spaces; Text/Textuality; Toponymy; Vision and Geography; Writing

Further Readings

focusing time and effort on evidence-based approaches to disease prevention and health promotion. Figure 1 shows a global view of HIV infection.

**Figure 1** A global view of HIV infection. According to UN estimates, about 33 million people worldwide were living with HIV/AIDS in 2007.

*Source*: Figure 2.2 from p. 33 in *2008 Report on the Global AIDS Epidemic* by UNAIDS. Copyright © Joint United Nations Programme on HIV/AIDS (UNAIDS) 2008. Reprinted with permission from UNAIDS.

### Historical Foundations

Although the acknowledgment of the relationship between environment and health can be traced back to 400 BC, when Hippocrates wrote *On Airs, Waters, and Places*, the first high-profile, and best-known, study of disease geography is likely that undertaken by John Snow. There was an epidemic of cholera in London during the year 1854. Snow, who hypothesized that cholera was a water-borne disease, plotted cases of cholera in the London neighborhood of Soho on a dot map and identified a water pump on Broad Street as the likely source of the epidemic. The handle was removed from the pump, disabling it, and the epidemic abated. Snow’s use of dot maps in identifying the source of the epidemic was novel and is the earliest case of the use of disease mapping in addressing a health problem.

### Causes of Spatial Patterning

There are many reasons for the patterns of disease that we observe. In Snow’s case, the reason for the spatial patterning of cholera cases in Soho was the location of a pump making contaminated water available to the local population. It is clear that the quality of the physical environment in which people live is of great importance in affecting rates of disease. The physical environment includes both the natural environment and the built environment. Commonly explored features of the physical environment, which can affect spatial patterns of disease, include air and water quality and sources of pollution that may be point sources (such as facilities releasing toxics into the air), nonpoint sources (such as agricultural runoff), and mobile sources (such as automobile exhaust). Climate is a factor of growing importance due to global climate change, which has the potential to affect the locations where disease vectors, such as mosquitoes that carry malaria, can survive. Built environment considerations range from the quality of indoor air affecting asthma rates to the walkability and safety of neighborhoods affecting rates of obesity.
In addition, spatial patterns of underlying social and economic conditions or population demographics can have a significant effect on disease burdens. For example, a poor neighborhood may be more likely to have higher rates of lead poisoning due to old or badly maintained housing stock containing lead paint, or higher rates of respiratory diseases due to a local, polluting facility. A working-class neighborhood may be more likely to suffer higher rates of musculoskeletal injury, due to the occupational duties of many residents. A neighborhood inhabited by people of a certain race, ethnicity, or origin could have higher rates of a particular disease associated with genetic or cultural factors. Finally, poverty can be associated with lack of financial access to health care and could increase rates of late-stage cancer diagnosis, for example, as people cannot afford the preventive care necessary to identify cancers in the earlier stages, waiting instead for symptoms to arise. Age is a factor likely to affect burdens of many diseases.

Population density can also affect spatial patterns of disease as individuals in rural areas may not have the same access to health care facilities as those in more populated places. Likewise, occupations associated with rural areas—such as farming—are known to have an increased risk of certain types of injuries. There is a growing interest in studying inequality and injustice aspects of differential burdens of disease associated with specific population groups.

In considering infectious diseases, the rates and methods of transmission must be considered. The length of time that a person can be infected but show no symptoms, the length of time an individual can be infected, and the characteristics of individuals that make them susceptible to infection all come in to play. Disease patterns may observably diffuse across space from a certain location of origin or may spread in a cascade, as individuals infected early on may travel to new locations, resulting in additional centers from which diffusion may occur.

**Developments in the Field**

Since the time of Snow, much has occurred in the way of developments in computer technology and statistics and the availability of geographically referenced health data, making disease geography a growing focus of study. The most important development may well be the birth of the field of geographic information science (GIScience) and the accompanying software programs that fall under the heading of geographic information systems (GIS). GIS are computer-based systems for integrating and analyzing spatially referenced data. These developments both increased the sophistication of the underlying theory guiding the study of disease geography and made it possible for researchers to undertake complex analyses without requiring large amounts of computation time. GIS are organized to allow the integration of different types of geographical information (such as data associated with points, lines, polygons and field presentations, such as rasters) and to make readily available a number of cartographic and statistical operations that can be performed on these data. In tandem with developments in GISc and GIS were developments in statistics—specifically spatial statistics—that make possible statistical modeling of risk, detection of clusters, hypothesis testing, simulation, and association studies.

The application of GIS and spatial statistics to health questions would not be possible without sufficient georeferencing of health data. In recent years, it has become more commonplace for disease registries and public health authorities to attach geographic codes—or geocodes—to health data, thus enabling the consideration of space in epidemiological and other analyses. Geocodes can include latitude-longitude of residence, ZIP code, or census tract information, and the available geocode often has significant implications on the type of analyses that can be performed. Along with developments in geocoding, there has been increased concern with the confidentiality implications of the availability of finely geocoded health data. Because of the ability to reverse geocode health records (i.e., determine the address of a residence associated with a particular latitude-longitude coordinate attached to a disease record), data holders must be cautious in releasing geocoded data.

A major regulation affecting the use of geocoded health data is the Health Insurance Portability and Accountability Act (HIPAA) of 1996. This act, implemented by the HIPAA Privacy Rule, prohibits covered entities (such as disease
registries) from disclosing private health information about an individual without his or her written, informed consent. Identifiers that make health information individually identifiable do include geocodes, making the use of these data highly regulated. The rule is qualified by exceptions, including mandatory reporting of diseases to public health authorities and the conduct of research by approved parties.

**Current Methods**

Current methods of studying the geography of disease include mapping, statistical analysis, simulation, and qualitative investigation. The most basic of exploratory methods is likely the mapping of a selected health statistic (e.g., a mortality rate) by a selected geographic unit, such as the county or census tract. These choropleth maps make use of discrete, bounded units and have been criticized for several reasons. Disease burdens do not adhere to borders of administrative units, and thus, the presentation of disease maps based on these units is misleading. Some have described the maps that result from this method as smoothed maps of disease data with the bounded areas chosen acting as spatial filters of different shapes and sizes—making the approach hard to defend.

In addition, mapping data in this fashion creates a surface populated by areas for which statistics have been calculated based on differing amounts of spatial support, making some statistics more reliable than others. In this case, for example, counties containing significant urban areas will have more stable rates than very rural counties due to higher population densities. For some units, such as census tracts, the area covered by a unit will be much larger in rural areas than in urban areas, giving visual focus to rural areas, where fewer people live. Also, mapping using discrete, bounded units masks all the geographical variation that occurs at a level below that of the unit chosen, giving the impression that the rate is constant throughout the entire areal unit. A final concern is that of the modifiable areal unit problem (MAUP), as the size and shape of the geographical units chosen to map disease burdens affect the overall spatial pattern that can be observed.

Approaches developed to improve the mapping of disease burdens include methods of interpolation, filtering, and smoothing. Some have smoothed data for discrete units by calculating a new rate for each unit—that of the average of the unit in question combined with all units that border it. A more recent approach, called *headbanging*, has used the median value instead of the mean to more carefully smooth rates. Still others have implemented methods of Bayesian adjustment, which shrink extreme values toward a central value to limit extremely high or low rates.

Other approaches have begun to disregard borders entirely, creating continuously varying surfaces. These approaches include geostatistical techniques such as Kriging, as well as density estimation and spatial filtering. In the most basic sense, Kriging smooths data based on the spatial dependence inherent in the data itself and was actually developed to enhance mineral exploration and mining. Density estimation and spatial filtering approaches make use of a grid and use filters of a selected shape to pull in data from neighboring areas to calculate a statistical rate for each grid point. These filters can be associated with functions whereby observations closer to the center of the filter are more heavily weighted than those at the edges of the filter. They can also vary in size. Whereas fixed filter spatial filtering would set the size of the filter to be constant (e.g., a filter with a diameter of 5 miles), adaptive spatial filtering would set a threshold for the denominator (e.g., pull in observations until a certain population threshold has been met), thus allowing the filter to expand variably across the surface to create a surface with more statistically stable rates throughout.

In addition to exploratory analyses such as mapping are analyses that seek to identify statistically significant rates of a disease. These analyses often make use of cluster detection statistics, which compare, for example, the rate observed at a particular location with what the expected rate would be, given, for example, a standard rate applied to a local population. A popular cluster detection statistic is SaTScan, which uses a circular or elliptical window to scan a data set through space and time to identify significant clusters. Similarly, statistics can be used to test hypotheses using spatial data. Because of the advent of Markov Chain Monte Carlo (MCMC) statistics, which allow the repeated simulation of
DISEASE, GEOGRAPHY OF

an expected spatial distribution, given an a priori statistical distribution or hypothesis, researchers can create what are known as *p*-value maps, which indicate the reliability of rates calculated across a study area using a disease mapping method.

Spatial autocorrelation, the likelihood of things closer in space being more similar, is a fundamental consideration in spatial analysis. Because of spatial autocorrelation, methods such as *p*-value map calculation must be used with caution as spatial dependence in the data violates the assumption of independent observations, which underlies many statistical tests. In the case of *p*-value maps, this means that there is a problem with multiple testing of the hypothesis. Statistics do exist to identify spatial autocorrelation in the data, including the global and local Moran’s *I* statistics and the *K*-function statistic.

Understanding the degree to which data are spatially dependent can inform approaches to modeling the data. For instance, residuals from a regression model that makes use of spatial units can be mapped to determine whether patterns exist that should be incorporated into the model. Models can then incorporate this autocorrelation, or other approaches can be taken, such as geographically weighted regression, which allows regression parameters to vary over space.

Simulation, including MCMC approaches, has been increasingly used with health data. Models such as the susceptible-infected-recovered (SIR) model for disease transmission can be implemented using simulation models to estimate the way in which an epidemic could spread. More recently, agent-based models have been used to simulate disease transmission through a set of probabilities associated with each agent.

Finally, in addition to the quantitative approaches described above is an entire set of qualitative research methods, such as interviews, focus groups, participant observation, and more structured exercises such as concept mapping, diagramming, and photovoice, which can be used to explore the mechanisms and causal pathways by which geography influences disease burdens. Theoretical constructs such as neighborhood, community, place, and social space can be used to guide these activities.

---

**Conclusion**

Disease geography is a growing field of study that has benefited from recent developments, including the advent of GIS and the increasing availability of geocoded data. New approaches to mapping and analyzing geocoded health data are developed frequently, and the importance of the “where” in understanding diseases makes it certain that the field will continue to develop. The geography of disease is of interest at many different geographical scales, from the impacts of global climate change and modern transportation on the global spread of disease to the impact of the choice of location for a health care facility in a small town.

Future directions for the study of geography of disease include the increasing involvement of the public and affected communities in exploring geographical data. Web-based map servers and software packages are increasingly moving the field toward a more democratic use of technology and information. New questions regarding the importance of area level influences—such as the level of deprivation in a census tract—are leading the field toward consideration of multilevel models, which consider both individual and contextual risk factors for disease. Spatially aware approaches to multilevel modeling are also under development. There are many opportunities for new approaches to studying the geography of disease, and these approaches have the potential to yield much new information, as well as new approaches to combating diseases.

*Kirsten M. M. Beyer*

*See also* Anthropogenic Climate Change; Cancer, Geography of; Carcinogens; Cholera, Geography of; Diffusion; Environmental Justice; Environmental Racism; GIS in Health Research and Health Care; Global Environmental Change; Global Warming; Health and Health Care, Geography of; HIV/AIDS, Geography of; Inequality and Geography; Malaria, Geography of; Medical Geography; Public Participation GIS

---

**Further Readings**

Distance decay is a major cornerstone of spatial-temporal analysis and, hence, geography. Essentially, it states that the nearer two locales are, the greater is the expected attraction/interaction between the two and, conversely, the greater the distance is between two locales, the less the expected interaction. This seemingly commonsense notion has been restated as a formulation of spatial autocorrelation by the geographer Waldo Tobler. It has come to be known as Tobler’s “first law of geography,” which states that “everything is related to everything else, but near things are more related than distant things.”

In its simplest form, the distance decay function can be portrayed graphically as a concave curve descending from the y axis, representing the intensity of a phenomenon, and extending along the x axis, representing the distance from some origin. Examples of this phenomenon can be found in virtually all subfields of geography and at all scales for a wide range of travel interactions. The universality of distance decay necessarily limits the discussion here to travel and human geography. In human geography, the distance decay effect has been noted in the movement of international tourists; travel to local coffee shops in Japan; use of mental health clinics in New York City and Baltimore; behavior of a serial killer in Baton Rouge, Louisiana; and forest land use patterns in the northeastern Philippines. An example of the distance decay function for a series of everyday activities is presented in Figure 1.

Distance decay functions are quantitatively rooted in the family of gravity models based on Isaac Newton’s fundamental law of attraction. The term distance decay characterizes how the attraction between two bodies decreases as the distance between them increases. The distance decay concept can be traced to the principle of least action, first espoused by Pierre Louis Moreau de Maupertuis (1698–1759), who first gave the notion verbal expression:

The laws of movement and of rest deduced from this principle being precisely the same as those observed in nature, we can admire the application of it to all phenomena. The movement of animals, the vegetative growth of plants... are only its consequences; and the spectacle of the universe becomes so much the grander, so much more beautiful, the worthier of its Author, when one knows that a small number of laws, most wisely established, suffice for all movements. (1744, p. 417)

It is difficult for any single discipline to “claim” Maupertuis for he wrote in physics, mathematics, metaphysics, and biology. He was the force behind expeditions to Lapland and Peru, traveling to the former to measure the difference in Earth’s radius and demonstrate that Earth was an oblate spheroid rather than ovate (egg shaped).

The concept of distance decay formulated by Maupertuis, although somewhat deterministic today, does capture much of the essence of mechanics. In applying it to physics, Maupertuis suggested that the quantity to be minimized was the product of the duration (time) of movement within a system and the vis viva (living force), twice what we now call the kinetic energy of the system.
Ravenstein (1885, 1889) is generally credited with the incorporation of distance decay theories into migration and spatial interaction studies. He noted in his “laws of migration,” for example, “The more distance from the fountainhead which feeds them (migrants), the less swiftly do these currents flow” (1885, p. 191). And “[We have] proved that the great body of our migrants only proceed a short distance” (p. 198). Furthermore, “Migrants enumerated in a . . . center of absorption will . . . grow less with the distance proportionally” (p. 199). Finally, “The process of dispersion is the inverse of that of absorption, and exhibits similar features” (p. 199).

Ravenstein’s distance decay observations pertaining to migration are frequently cited as among the earliest observations pertaining to human behavior. However, Edward Jarvis observed, documented, described, and, in fact, explained the impact of distance on travel behavior in his assessment of the admission rates to “insane asylums” of the mid 19th century. He

Figure 1  Distance decay curves for single-occupant auto trips. Automobile trips, like other forms of spatial interaction, exhibit distance decay as the friction of distance makes longer trips more costly.

Source: Figure 7, p. 19 from Access to Destinations: How Close Is Close Enough? Estimating Accurate Distance Decay Functions for Multiple Modes and Different Purposes by Michael Iacono, Kevin Krizek, and Ahmed El-Geneidy, May 2008. Published by Minnesota Department of Transportation.
was writing at a time when there was concern in the United States and Europe “that insanity (had) increased in prevalence of late years, to an alarming extent, and that the number of lunatics, when compared with the population, is continually on the increase” (Jarvis, 1852, quoted in Shannon & Bashshur, 1986). In these articles describing what today is known as “Jarvis’s Law,” he demonstrated that the number of patient admissions (amount of use) to mental hospitals of the period was inversely proportional to the distance at which the patients lived from that asylum (service) (Shannon & Bashshur, 1986). In 1852, he documented the phenomenon in an article with insane asylum admissions data from New York and Massachusetts to Kentucky and Ohio. Beyond merely documenting the distance decay phenomenon, however, Jarvis provided an explanation that must be considered cogent 170 years later. Among the factors he noted as possibly contributing to the distance decay effect were (a) cost of travel, (b) unwillingness of relatives to be separated from loved ones, and (c) lack of accurate information pertaining to the benefits of asylum care with increasing distance from the institution.

The distance decay principle is a cornerstone of spatial analysis. As such it is fundamental to classic models and theories of geography as well as current analyses in virtually every subject matter addressed by geographers.

Gary W. Shannon

See also Gravity Model; Migration; Mobility; Spatial Autocorrelation; Spatial Interaction Models; Tobler, Waldo

Further Readings


**Distributed Computing**

Distributed computing is an important research domain in computer science and information technology. It plays a significant role in the development of next-generation geographic
information systems (GIS), including Web GIS and mobile GIS. The major goal of distributed computing is to provide an effective, open, and dynamic computing framework to facilitate via networks the integration of data, applications, and services physically located in different places. There are several distributed computing frameworks available today, such as Microsoft .NET, the JAVA platform, and Common Object Request Broker Architecture. For many GIS applications, these frameworks provide sharable and interoperable software environments for developing online, distributed geographic information services (GI Services). The deployment of high-level distributed GI Services is built on low-level distributed computing technologies. Distributed computing technologies can facilitate the interoperability and federation of GIS computing tasks.

The term distributed reflects the fact that the hardware and software components of distributed computing systems are physically distributed in different computers, which are connected via the Internet or other types of networks. Interoperability is the key issue for the establishment of distributed computing frameworks because distributed hardware machines, programming languages, operating systems and other online resources may vary drastically. To provide online information services effectively, most distributed computing applications use open and interoperable computing environments and protocols (such as TCP/IP, File Transfer Protocol, Remote Procedure Calls, and Z39.50) and distributed programming languages (Java, JavaScript, Python, or C-sharp) to connect multiple machines and servers together.

**History of Distributed Computing**

Along with the rapid progress of network technology, distributed computing technologies have been widely used in the computer industry since the 1980s. Suggested by Orfali, Harkey, and Edwards in 1996, the development of distributed computing systems can be categorized into four major stages: (1) stand-alone file servers, (2) generic database servers, (3) distributed database servers and file servers, and (4) distributed component object servers. The following paragraphs discuss the four stages by their definitions, network features, and system structures.

**Stand-Alone File Servers**

A file server is a device that stores and distributes digital files to every user on a local area network (LAN). It allows everyone on the network to access files from a centralized storage space, on one server. A file server directs movement of digital files and data on a multiuser communication network. From a network management perspective, file servers usually handle a huge amount of transactions, which might become a significant bottleneck in a LAN. Different file servers such as Windows servers and Unix servers have their own protocols and file formats, which may not be compatible with others.

**Generic Database Servers**

A generic database server is a stand-alone computer that sends out database data to users on a LAN. With a database server, the server only sends out the requested results from a database rather than send the whole database to a user’s workstation. Thus, a database server incurs less network traffic than a file server. Database servers are more flexible than file server systems, especially on the client side. Multiple users can easily establish new client-side applications to access the same database server from different operating systems. However, the server-side applications are fixed in most cases. It is very difficult to combine heterogeneous databases in one single server.

**Distributed Database Servers and File Servers**

The main functions and capabilities of distributed database servers mimic generic database servers, but the physical locations of databases are distributed across a network. As with the architecture of distributed databases, distributed file servers appear as a single logical file server but are physically distributed in different places. Distributed file servers are designed for file sharing instead of database access. Distributed file servers can provide users with a virtual integration of distributed file servers on a LAN. An
example of this is the active directory services in the architecture of Microsoft Windows 2000 and XP. Both generic and distributed database/file server systems basically follow the traditional client-server architecture. There are several potential problems with the traditional client-server architecture because it cannot provide rich transaction processing and rich data management or handle overly complex queries. For example, if a traditional database server receives requests from 500 client-side applications at the same time, the server’s operating system may hang or crash. Without a transaction control function, traditional database architecture is not appropriate for complex GIS applications or Web GIS services.

### Distributed Component Object Servers

Distributed component object servers are advanced client-server systems that can handle complex transactions and requests from heterogeneous systems. Distributed component technology, including Java applets and ActiveX controls, adopts the concepts of object-oriented modeling and distributed computing environment. Currently, both academic and industrial studies of distributed systems are focusing on distributed components in open environments that can provide new capabilities for the next generation client/server architecture. Comparing the distributed database/file servers, the main advantage of distributed component object servers is the interoperability, reusability, and flexibility for cross-platform applications.

In summary, traditional GIS are closed, platform-dependent, centralized systems, incorporating interfaces, programs, and data in a single computer. Distributed computing technologies provide a new direction for GIS to become more flexible, interoperable, and platform independent. Many wireless mobile GIS applications and Web GIS services are adopting the concepts and technologies of distributed computing.

*Ming-Hsiang Tsou*

---

**Further Readings**


resource access is embedded in a broader study of human-resource relationships that also considers resource availability and the resource use practices of humans. For example, applied research and development planning regarding food security evaluates how the natural endowment of resources and production strategies influences food availability, how household incomes can be provided and improved to gain access to food resources, and how household practices and individual discrimination influence the distribution and quality of food for use. Access to resources and how it varies geographically relate to the need of regions and households to gain incomes to supplement local livelihoods and what sources of income can be made available to these populations. Regional studies of resource access are therefore linked to studies that examine the magnitude and problems associated with regional poverty at different geographic scales.

Garrett Hardin’s classic work on the “tragedy of the commons” emphasized that when resources are free or open to communities, overuse or exploitation will affect future generations. Access to “common” resources should therefore be managed under environmental regulations that dictate who gains access and for what purpose. The ways in which environmental (resource) regulations differ geographically and are directed at certain populations add a political dimension to the distribution of resource access. These regulations are defined by local “traditional” practices and/or by extralocal agendas designed to promote resource conservation and sustainability. Communal access and community management as directed by the local populations and their traditions are compared with the controls that may be imposed by authoritative structures. Resource access is therefore dictated more by boundaries and controls on the use of common resources than by environmental conditions and processes that influence distribution patterns. Moreover, the distribution of resource access is redefined by political structures that allocate ownership of common resources and/or authority over their use. Embedded in these political debates on resource access are questions about whose knowledge matters and who is in the better position to ensure the long-term sustainability of common property resources.

Geographic studies in historical ecology also document interesting relationships between the distribution of resource access and land cover dynamics. Technological changes and regional patterns of development can promote access to resources and consequent human activities. For example, historical change in the agricultural land use of the Central United States relates significantly to technological developments that provided access to the Ogallala aquifer, just as deforestation in the Amazon basin can be related to the development of roads and consequent settlement patterns. Resource access, while it is related to endowment, can be altered by technological development that provides and limits distributions among regions. Access to water resources have been grossly modified through the development of dams, irrigation canals, wells, and other access strategies that promote its distribution to some locations while limiting access to other locations. The distribution of resource access and its modification by technology and human resource practices can intensify differences in regional endowments in ways that lead to conflict across regional and state boundaries.

The distribution of resource access, according to geographers working in political ecology, relates to power relations in societies and among locations. For these studies, a special focus is placed on the relationship between access to resources and control over their management. Authoritative powers at both global and local scales can impose restrictions that limit access to common resources through the establishment of protected areas or other types of tenure arrangements, or they can redirect decision-making opportunities and management to designated populations. Power relations can be modified by redefining who has control over the management decisions that determine resource access and its use. Feminist research in political ecology focuses on how access and control differ within communities and households. Women may have access to resources, but the traditional practices of societies may limit ownership or dictate other forms of control over how these resources are managed and sustained. Accordingly, marginalized groups of people often experience a very different distribution of resource access in a region, and they possess little authority over how resources are
used and how distributions may be modified by human activities. Classic studies in feminist political ecology have examined women’s role in agricultural production, water collection, and fuelwood collection and critiqued how their roles are constrained by the distribution of resource access and control. Geographic studies on the distribution of resource access are necessarily integrative in their examination of resource endowments, human and resource relationships, and power relations within and among locations.

Kimberley E. Medley

See also Common Property Resource Management; Commons, Tragedy of the; Feminist Political Ecology; Land Tenure; Land Use Analysis; Resource Tenure

**Further Readings**


**DIVISION OF LABOR**

Perhaps Adam Smith’s most famous dictum is that “the division of labour is limited by the extent of the market.” By allowing individuals to specialize in the production of some goods and trade them for different goods (including arts and sciences) produced by other individuals, the division of labor was, in Smith’s eyes, the ultimate source of wealth and economic development. The bigger the market, the more numerous and varied the number of tasks and opportunities become.

For Smith, the benefits of the division of labor derive from three sources: (1) increased dexterity of a worker, (2) time saved from switching between tasks, and (3) increased likelihood of new inventions. The first two effects are today referred to as increasing returns. Unlike David Ricardo, Smith considered the inherent advantages of specialization as inferior to those acquired by learning and experience. The last benefit of the division of labor stresses the idea that individuals are more likely to come up with new ways to facilitate their work when they concentrate on certain tasks. Furthermore, Smith observes that some improvements have been developed by producers of machinery, while others were created by individuals who specialize into combining ideas from distant fields. In short, the division of labor yields not only static benefits through better organization of production but also new innovations and improvements developed through time.

The division of labor has always had critics. In the early days of industrialization, it was feared that it would result in ever duller and more mechanistic work. Another recurring concern has been regional or social inequality. Yet the division of labor is also an argument on behalf of globalization since the expanding market allows for more division of labor and its attending benefits.

While the division of labor can be studied in various contexts such as households, three levels, following Allen Scott, are generally considered especially relevant for geographers. The first is the individual level, where individuals specialize in different tasks within a firm (*the technical division of labor*), in the process yielding a greater output for a firm, such as Smith’s classic example of a pin factory. The second division operates at the firm level, where firms in a market specialize in producing different goods and services (*the social division of labor*). Last, due to natural or other advantages, it is sometimes beneficial for regions (or firms in some geographical areas) to specialize in producing some goods and trading them with other firms and individuals located elsewhere (*the spatial division of labor*).

The geographical division of labor has two main dimensions. The first deals with how the divisions are achieved through space and how they are affected by geographical factors. Second, the division of labor being “limited by the extent
of the market,” it is interesting to study the spatial borders of the market and their evolution over time, as well as trade and other linkages between locations.

As Richard Baldwin and Philippe Martin have noted, there have been two major phases in this respect in the past two centuries. During the so-called first unbundling (1820–1914), the development of transportation technologies allowed the spatial separation of the production and consumption of many goods. As a result, the division of labor became less dependent on the size of the local demand. The second unbundling (1960 to the present), however, was characterized by the separation of different stages of production. In addition to transportation technologies, advanced communication technologies were helpful in coordinating the production of goods when different inputs and some supporting services became increasingly geographically dispersed. These impacts of new technologies, themselves products of the division, underline the dynamic character of the division of labor.

A deepening division of labor increases the interdependence of individuals and societies, which often creates negative reactions against increased dependence on foreign goods and services. As Friedrich Hayek noted, however, the division of labor is analogous and strongly connected to the division of knowledge. People know little about the trade of others and will rarely be as good at doing these tasks themselves but can nonetheless benefit from others’ expertise through market exchange. The market mechanism—and especially the price system—is therefore crucial for coordinating interdependent activities when the complexity of the economic system increases due to specialization. Besides the extent of the market, the division of labor is thus also limited by coordination costs and the amount of general knowledge. While the division of labor provides opportunities for further developments, it also generates new challenges and links economies more closely together, now on a global scale.

Samuli Leppälä and Pierre Desrochers

Further Readings


DOMESTICATION CENTERS

See Centers of Domestication

DOMESTICATION OF ANIMALS

For most of human existence, people lived as seasonally mobile gatherers, hunters, and fishers. A profound change in the human relationship with nature occurred when people established control over the reproduction and evolution of owned populations of other species, which became domesticated plants and animals. The early integration of agriculture and animal husbandry has been called the “Neolithic Revolution” because of its epochal transformation of the nature-society relationship. The domestication of animals is a major part of this transformation, and geographers have investigated it from their discipline’s earliest formalization, in the 19th century. Early geographers who wrote
DOMESTICATION OF ANIMALS

The domestication of animals included Alexander von Humboldt, Friedrich Ratzel, and Eduard Hahn and, in the 20th century, Carl Sauer, Erich Isaac, Frederick and Elizabeth Simoons, and Carl Johannessen.

Archaeologists, ecologists, and zoologists share this interest, but the geographers’ approach differs from those of others in three distinctive ways. First, the geographic tradition in animal domestication studies uniquely stresses the role of ritual and religion in motivating people to take on the work and risk of animal husbandry.

Second, many geographers assert a three-stage scheme of cultural ecological change, from migratory gathering-hunting-fishing societies, through settling down in villages securely supported by fishing and successful experiments with plant propagation, to the domestication of animals, all in stable environments not strongly affected by postglacial changes. They argue that economic necessity could not have motivated such well-fed folk to domesticate animals, and so these sometimes dangerous endeavors were driven by the desire to get along with the supernatural.

A third focus of the work of many geographers is diffusion(ism). This attempts to pin down the original “hearth” of domestication and then trace out the paths along which the animals themselves or the idea of their active management might have diffused.

The evolution of animals under human control eventually led to significant genetic and phenotypic divergences from wild populations. Some phenotypic changes allow advanced animal domestication to be recognized in faunal remains. Earlier stages can be detected in changing age-sex structures. The archaeological and paleoenvironmental...
DOMESTICATION OF ANIMALS

Almost certainly the animal that has been domesticated for the longest time is the dog—actually a wolf, *Canis lupus* (*C. lupus familiaris*). The oldest archaeological remains of domesticated dogs go back about 14,000 cal. (calibrated) BP (years before the “present” of 1950) in Germany and 12,000 cal. BP in Israel, while analysis of mitochondrial DNA (mtDNA) in modern dogs suggests descent from three females in east Asia some 15,000 years ago (roughly the end of the last Pleistocene glaciation). Changes in the age-sex structure of gazelles (*Gazella gazella*) in the Levant imply that the Late Natufian culture of 11,500 to 12,800 cal. BP initiated their domestication during the cooling and drying of the Younger Dryas.

A second wave of animal domestications followed quickly after the dog’s (and gazelle’s), in the 11,000 to 10,000 cal. BP time frame in the early Holocene Hypsithermal warming. These have long been understood to include sheep, *Ovis orientalis aries* (Northern Iraq), and goats, *Capra aegagrus hircus* (Western Iran), but recent analyses have pushed pig (*Sus scrofa domesticus*) domestication back to the same time in southeastern Turkey and Northern Syria. Recent DNA analyses suggest that cattle, *Bos primigenius taurus*, long considered a much later domesticate, were domesticated in that same 11,000- to 10,000-BP window in northeast Africa. Chickens, *Gallus gallus domesticus*, may also go back to this time frame or slightly later, according to mtDNA analyses, with origins in southeast and South Asia. Archaeological finds on Cyprus put the domestication of cats (*Felis sylvestris catus*) at 9,500 cal. BP.

This second wave shows animal domestication underway in several specific locations in two broad regions, the southwest Asia/northeast Africa region and the southeast/South Asia region, loosely supporting the notion of a hearth. For decades, however, geographers and archaeologists dismissed significant Pleistocene-Holocene environmental change in the Near East as relevant to early domestications. Physical geographers increasingly document the magnitude and, often, abruptness of environmental changes at that transition, both in the Near East and southeast/South Asia, as well as globally. They are an important context to domestication processes in the 14,500- to 10,000-BP time frame.

Within these two broad regions, particular animal domestications evolved in specific locations that both overlapped and differed from those in which plant domestications were going on at the same time (not long before, as the geographers’ three-stage models had assumed). Animal and plant domestications were integrated into agricultural and/or horticultural complexes through rapid crisscross diffusion from one place to another within these regions. These domestications then exploded outward with rapidly expanding human populations and with the ensuing exchange of resources and ideas.

A third wave of animal domestications followed much later, from 7,000 to 4,500 cal. BP. This involved the New World and extensive portions of the Old World. Species domesticated then included the alpaca (*Vicugna pacos*), llama (*Lama glama*), and guinea pig (*Cavia porcellus*) in Peru; the pigeon (*Columba livia domestica*) in Mesopotamia and Assyria; the horse (*Equus ferus caballus*) in the Ukraine, Southern Russia, and Western Kazakhstan; the ass (*E. africanus asinus*) in northeast Africa and possibly Somalia; the Bactrian (*Camelus ferus bactrianus*) camel in northwest China, Turkmenia, and Eastern Iran; the dromedary (*Camelus dromedarius*) in the Arabian Peninsula; several bovines in India, Southern Iraq, southeast Asia, and Tibet; and the silkworm (*Bombyx mandarina mori*) in China.

Newer domestications have continued all over the world to the present. Some of these are utilitarian, for meat, fur, honey, working, or laboratory use (turkeys and rabbits, mink, bees, cormorants, and rats); some are experimental, with an eye toward utilitarian value (eland, fallow deer, ostrich, and foxes), while others are for pet keeping (parakeets, koi, “fancy rats”).

Christine M. Rodrigue

See also Animal Geographies; Centers of Domestication; Climate Change; Diamond, Jared; Domestication of Plants; Humboldt, Alexander von; Hunting and Gathering; Nomadic Herding; Ratzel, Friedrich; Sauer, Carl
DOMESTICATION OF PLANTS

Domestication is the evolution and creation of a mutualistic, interdependent relationship between humans and wild plants (and animals). During the domestication process, wild plants become selected for and adapted to human cultivation, use, and consumption. Domesticated plants are an inherent part of agriculture, and their appearance, evolution, and dissemination throughout the world is closely intertwined with the origin, spread, and current dominance of agriculture. The study of the domestication process is therefore in large part also the study of the origin and evolution of agriculture. Major lines of inquiry are determining the actual locations of agricultural origins and their characteristics, the causes for the adoption of agriculture, the effects of selection on plants during domestication (leading to the so-called domestication syndrome), and the consequences of the adoption of agriculture on a worldwide scale.

Domestication as a Genetic and Evolutionary Process

During the process of domestication, three sets of factors are of paramount importance to determine the degree, speed, and ultimate success of domestication. These include biological characteristics such as the length of the plant’s life cycle and its reproduction method, environmental conditions such as the climate, and human cultural traits such as knowledge of the surrounding environment and toolmaking techniques. During domestication, evolutionary processes come into play. For example, naturally occurring or de novo mutations are selected, consciously or unconsciously, by humans. Mutant traits selected by humans increase yield under cultivation, enhance the flavor of the harvest (or reduce its toxicity), or improve the quality of the product. Demographic processes such as small sample size also influence the process of domestication.

The domestication selection process has led to profound changes in the morphology, biochemistry, and physiology of plants. These domestication traits are similar in widely different species and are collectively called the “domestication syndrome.” For grain and seed crops, the two most important traits of the syndrome are the loss of seed dormancy and dispersal. Both traits are essential in natural environments but are undesirable under cultivation. In addition, domesticated plants tend to be smaller than their wild progenitors, with fewer and shorter branches. However, the harvested parts, such as leaves, seeds, and fruits, tend to be larger and show more diversity in shape and color than do their wild progenitors. This diversity is a direct consequence of selection by humans, who favor novelty. Selection during domestication has been so marked in several crops that domesticates have lost the ability to survive without human intervention. Humans and their crops have, therefore, established an obligate mutualism, both depending on each other for survival.

Geographic Centers of Domestication

Agriculture originated independently in several regions of the world in so-called centers of agricultural origins or domestication, which are generally located on either side of the equator, between latitudes 35° N and 35° S. Many of these centers are located in bioregions with an alternation of dry and wet seasons, such as the Mediterranean, savanna, and tropical deciduous forest biomes. This alternation may have stimulated the transition from gathering to agriculture. Domestication centers are generally located in areas of high biodiversity that are often characterized by a heterogeneous topography. Biodiverse regions provided opportunities for the first farmers to experiment with a wide range of plants and to identify species that were more amenable to domestication. Figure 1 shows the approximate location and boundaries of six major centers of domestication and agricultural origins.

In each of these centers of agricultural origin, a similar complement of crops has been domesticated that address basic human needs or activities.

Further Readings

Crops can be categorized as sources of carbohydrates, protein, oil, fibers, and stimulants. Each center has developed a complex of crops with one representative from each of these categories. In addition, some of these crops offer complementarity from an agronomic and human nutrition standpoint. For example, legumes such as beans and peas complement cereals by providing nitrogen to cereal plants in the field and by adding protein to the human diet. Thus, domestication should be seen not only as a process affecting individual crops but also as a process leading to more complex agricultural systems consisting of multiple crops grown in association or rotation.

Figure 1  Approximate location and boundaries of the six major centers of domestication and agricultural origins. For each of these centers, a sample of crops domesticated in them is shown.


Timing of Domestication

A striking feature of these multiple origins of domestication is their near-simultaneity around the world. Archaeobotanical research has identified remains of domesticated plants that are as old as 10,000 years in several of these domestication centers. Around the same time—some 15,000 to 10,000 years ago—global climate change associated with the end of the last Ice Age raised the temperature and caused changes in rainfall patterns and distribution. This situation, and other factors, may have led to an imbalance between the supply and demand of food and other crop products. Humans responded by supplementing
their food intake obtained from gathering with food derived from cultivation. The conversion to agriculture may have been helped further by the more stable climate appearing at that time.

**Dispersal of Agriculture and the Columbian Exchange**

In its strictest definition, domestication pertains to the process that took place in the incipient stages of the invention and adoption of agriculture. However, subsequent events are also part of domestication in a broader sense. Because agriculture provides a higher production per unit area, it outcompeted gathering and spread to all but the most extreme environments of the globe. Through agriculture, crops began to spread beyond their centers of origin as they were selected to adapt to new environments. The most important dispersal event was the Great American (Columbian) Exchange, which took place after the first European contact with the American continent in 1492. In the following years, crops from the Old and New Worlds were exchanged, forever changing the face of agriculture and human subsistence on the planet. Through this exchange, the Italians discovered tomatoes; the Irish, the potato; and Thailand, the chili pepper—formerly exotic ingredients that now form the backbone of these cuisines.

**Current Importance of Domestication Centers**

Modern agriculture is characterized by many crops being cultivated far outside their initial domestication center. However, the diversity of a crop is often greatest in its respective center of domestication due to environmental and cultural factors. While efforts at preserving crop diversity in seed banks (ex situ conservation) have several advantages, maintaining crop diversity on-farm (in situ conservation) has many benefits in addition to the maintenance of biodiversity, such as improved human health and economic status. A mechanism to reward the farmers, the stewards of this diversity, must be developed to ensure that future generations can use this resource to develop improved crop varieties in the face of looming challenges—large human population increases and the global climate change our world is witnessing.

*Paul Gepts, Matthew B. Hufford, and Kraig H. Kraft*

See also Agriculture, Preindustrial; Agrobiodiversity; Centers of Domestication; Crop Genetic Diversity; Diamond, Jared; Domestication of Animals; Food, Geography of; Great American Exchange; Hunting and Gathering

**Further Readings**


**DOMINO THEORY**

The domino theory was an important component of the U.S. Cold War geopolitical model of the world, which divided the world into three parts: the First World (the West), the Second World (communist states), and the Third World (colonies and newly independent states). Proponents of the domino theory argued that the loss of one
country in the First or Third World to communist control would trigger the loss of its neighboring countries to communism.

While the origins of the concept have been traced to early in the 20th century, the domino theory emerged in full force in U.S. foreign policy circles in the aftermath of World War II. The concept was advanced in 1947 both by the former U.S. ambassador to the Soviet Union, William Bullitt, who warned of communism advancing out of the USSR and sweeping over the world, and President Truman’s Secretary of State, Dean Acheson, who warned that a communist victory in Greece would lead to communism sweeping through Africa and Western Europe.

The domino analogy was first used in 1954 by U.S. Admiral Arthur Radford to justify his call for using nuclear weapons in support of French attempts to hold Indochina. President Eisenhower used the term *domino* that year to argue for U.S. intervention, arguing that “the loss of Indochina will cause the fall of southeast Asia like a set of dominoes” (quoted in O’Sullivan, 1982, p. 58). The domino theory reached its highpoint in U.S. foreign policy circles in the 1960s, when it helped guide the Kennedy and Johnson administrations’ escalation of U.S. involvement in Vietnam. That escalation was based on a belief that if South Vietnam fell to the communists, then the rest of southeast Asia would fall in quick succession, with potential consequences beyond.

The domino theory continued to be used to convince the American public of the importance of containing communism into the 1980s. For example, the Reagan administration sold its support of the contra rebels fighting the socialist Sandinista government in Nicaragua to the American public in part by arguing that the United States had to defeat the Sandinistas because if left unchecked, communism would spread throughout Central America and Mexico, bringing it to the Texas-Mexico border.

Throughout the Cold War, the domino theory suggested that one communist victory in some far-flung part of the world would lead to a series of victories in neighboring countries. Through this metaphor, distant events could be linked back home. Thus, intervention in southeast Asia in the 1960s and in Central America in the 1980s was portrayed as necessary because once countries in southeast Asia and Central America “fell,” then, like a row of falling dominoes, communist forces would ultimately continue until they reached American soil.

However, over time, the domino theory has been discredited, both empirically and theoretically. Attempts to demonstrate the domino theory in action have fallen short, and critics have argued that the domino theory’s mechanistic approach to the fall of countries ignores the vast differences and complexities among them. Despite the fact that the domino theory has been thoroughly discredited, it is still invoked by some today, for example, to predict the spread of terrorist activity.

Jonathan Leib

See also Cold War, Geography of; Geopolitics; Political Geography

Further Readings


Dot density maps, or dot maps, portray the geographic distribution of discrete phenomena using an arrangement of identical point symbols, most
commonly dots. The dot density technique dates to at least the 19th century and is today accepted as one of the primary techniques for representing geographic patterns. Dot density maps are particularly useful for understanding global distribution of the mapped phenomenon and comparing relative densities of different regions on the map. Dot density maps are also easy to understand, requiring little cognitive effort from the map reader when compared with isoline maps. However, retrieval of specific information from dot density maps is difficult as map users find manual counting of dots tedious and tend to underestimate dot totals as the density increases.

### Types of Dot Density Maps

There are two kinds of dot density maps: one-to-one maps and one-to-many maps. In one-to-one dot density maps, each point on the map corresponds to a single incidence of the mapped phenomenon. One-to-one dot density maps are general-reference maps, symbolizing spatial location only. Because of this, care should be taken to ensure that a dot is accurately located on the map. Examples of data sets ideal for one-to-one dot density mapping include the major cities in Europe or locations of recent earthquakes along the Pacific Rim.

Although one-to-one dot density maps are more common in practice, the term dot density map typically refers to one-to-many dot density maps. In one-to-many dot density maps, each point on the map represents a predetermined number of incidences of the mapped phenomenon, called the dot value. One-to-many dot density maps are thematic maps, symbolizing an aggregated variable atop a reference base map. Use of a one-to-many dot density map, rather than the one-to-one counterpart, is necessary when the only available data are aggregated to areal enumeration units or there are too many point incidences within the map extent for legible representation, necessitating aggregation by the cartographer. Examples of data sets ideal for one-to-many dot density mapping include the population of the United States and the number of dairy cows in Wisconsin, both aggregated by county.

#### Aggregated Data and One-to-Many Dot Density Maps

Not all aggregated data are appropriate for one-to-many dot density mapping. Alan MacEachren and David DiBiase developed a typology of aggregated data based on two characteristics of the mapped phenomenon depending on whether (1) the phenomenon occurs at discrete locations in space or exists continuously throughout the extent of the map or (2) the phenomenon changes abruptly at enumeration boundaries or varies smoothly throughout the extent of the map. Each data model is then paired with a recommended mapping technique. Figure 1 provides simple and effective guidance for determining the appropriate thematic map technique for representing aggregate data, given the characteristics of the mapped phenomenon. Only aggregated data of phenomena that exist discretely in space and vary smoothly across space should be mapped using the dot density technique. Only magnitude data should be displayed with dot density maps; derived values, such as averages, rates, and percentages, are theoretically continuous and should therefore be mapped with a choropleth.

#### One-to-Many Dot Density Map Design Considerations

There are four important design considerations for one-to-many dot density maps that affect the accuracy, clarity, and effectiveness of the image: (1) units of aggregation, (2) dot size, (3) dot value, and (4) dot placement. The size, shape, and distribution of enumeration units at which the data are aggregated can influence the mapped pattern greatly. Generally, the smaller the size and the more regular the shape and distribution of enumeration units, the more realistic the mapped pattern. When possible, enumeration units that are meaningful to the mapped phenomenon should be used over those that are arbitrary.
The dot size and dot value should be considered in tandem and together determine the number of dots on the map. An intermediate dot size should be chosen in most cases; dots that are too small produce an overly sparse dot pattern and convey an inappropriate amount of map accuracy, while dots that are too large produce excessively dense dot patterns, overwhelming subtleties in the distribution of the mapped phenomenon. A rounded, easily understood number should be chosen for the dot value for easy estimation. Selection of an appropriate dot size and dot value always requires some experimentation, but it can be made easier by using a visual tool called a nomograph. When an appropriate dot value and dot size are chosen, there should be two or three dots in the enumeration units with the smallest values, and the dots should just begin to coalesce in the enumeration units with the largest values.

The final important design consideration for one-to-many dot density maps is the method of

**Figure 1** A reference for matching mapping technique to aggregated data set based on the characteristics of the mapped phenomenon

DROUGHT RISK AND HAZARD

Drought is a complex concept with both physical and socially constructed components. The idea of a drought conjures a popular image of cracked brown earth, fallow agricultural fields, and suffering human communities. This realistic image confuses the physical drought hazard with the social impacts of drought in contexts with high drought risk. Drought does not lend itself to a single, universal definition because drought has different spatial and temporal characteristics in distinct climates and regions. Drought hazard refers to the physical phenomenon of a meteorological drought, generally defined as a departure from a long-term precipitation (e.g., rainfall or snowfall) average that results in an impact in a specific place. A drought is a temporary yet recurring phenomenon that occurs in all climates, but occurs with a greater frequency and more severe impacts under certain regional conditions than in others.

Unlike other hazards such as floods and tropical storms, a drought is gradual or “creeping.” Drought hazards are spatially vast and affect more people worldwide than do other “natural” disasters. Notable drought-prone regions of the world include Australia, Northern Mexico, northeast Brazil, Western Africa, and southwestern United States. With respect to drought, Africa experiences the greatest loss of life, Asia has the most people affected, and the West sustains the highest economic damages. Drought can lead to a massive loss of life, famine, desertification, deforestation, loss of livelihood, and migration.

Drought risk refers to the vulnerability associated with the exposure of an individual, a community, or a region to a drought hazard, as well as the capacity of that entity to adapt or respond in an effective way to the drought hazard. A drought hazard does not in itself cause a social disaster. Rather, socioeconomic attributes and policy choices determine whether a meteorological drought will create conditions of high drought risk. For example, in a poor country with a high proportion of subsistence farmers, a drought may imperil the country’s overall food security, while in a country with high levels of imported foods and widespread irrigation, a drought may have limited impacts. Drought hazard and risk are therefore separate yet related concepts.

Drought Types

Donald Wilhite, a geographer and former director of the U.S. National Drought Mitigation

Robert E. Roth

See also Cartography; Choropleth Maps; Dasymetric Maps; Isopleth Maps; MacEachren, Alan; Map Design

Further Readings


Center, defines four distinct types of drought: meteorological drought, hydrological drought, agricultural drought, and socioeconomic drought. At their root, all droughts originate as meteorological drought, defined as a deficiency in precipitation over a preestablished time period relative to an accepted “normal” longer-term historical average (e.g., 75% of normal precipitation over a 6-month period). Meteorological droughts in turn give rise to other identifiable types of drought. Agricultural drought refers to an insufficiency in soil water to support crop and forage growth; regional variations in crops and soil types help condition whether decreased precipitation develops into an agricultural drought. Hydrological drought occurs when there are insufficient surface and groundwater supplies to fulfill multiple and competing water demands, including irrigation, hydroelectric power, domestic use, and ecosystem management. Socioeconomic drought is at the nexus of the supply and demand for water or an economic good reliant on water (e.g., hydroelectric power). The supply is variable, based on precipitation and water availability, while the demand is conditioned by dynamic social factors such as demographic change and economic development trends. A precise definition of drought with regard to its type, timing, climate, and region can help in the development of targeted adaptive responses.

China, the Middle East, and India all experienced drought in prehistoric and historic times.
leading to dire consequences such as mass starvation, though these impacts were also conditioned by sociopolitical factors. In recent years, researchers have cited decade-scale variations in ocean-atmosphere interactions in the Pacific and Atlantic Ocean basins as the ultimate causes of severe sustained droughts. Paleoclimate studies confirm the influence of interannual-to-decadal ocean-atmosphere variability on extended drought episodes, more severe and extended than those in recorded memory, in North America, Asia, and the Mediterranean. North America tree ring drought reconstructions have provided evidence of multidecade “mega-droughts” from AD 900 to 1300, with devastating impacts, making it clear that drought contributed significantly to the dissolution of Pueblo Indian society in the American Southwest. Mega-droughts of similar magnitude may be a normal aspect of current and future climate regimes.

Measuring drought hazard is constrained by the lack of a common definition of drought. The departure of precipitation and reservoir levels from a threshold or decision-making criterion are common hydroclimatic drought indicators. Drought indices range from a straightforward percentage of average precipitation to supply-side indices, such as the Standardized Precipitation Index, to complex water balance formulae, such as the Palmer Standard Drought Indices, which factor in temperature. Other indices integrate information such as snowpack, streamflow, and reservoir storage or measure vegetation stress based on remote sensing.

The U.S. Drought Monitor and North American Drought Monitor combine multiple quantitative drought indicators with expert scientific assessment and impact reports to create snapshots of drought conditions each week and month, respectively. These drought indices subjectively associate drought status with drought impacts on resource management and agriculture. The U.S. National Drought Mitigation Center’s Drought Impact Reporter monitors and measures drought impacts through news media reports, volunteer reports, and economic losses.

Climate scientists seek to understand the relationships between climate and drought in multiple contexts. Drought occurs on many spatial and temporal scales, thus a variety of physical mechanisms have been associated with drought. On weekly-to-monthly timescales, diminished precipitation, late precipitation (e.g., delayed monsoon onset), sublimation of mountain snowpack, and heat waves may be the proximate causes of drought impacts on agriculture, streamflows, or wildlife habitat. On seasonal-to-decadal timescales, hemisphere-scale interactions between the ocean, atmosphere, and land are usually implicated. The El Niño-Southern Oscillation (ENSO) is strongly associated with interannual drought in the global tropics and subtropics and some midlatitude regions (e.g., Western North America). ENSO effects on seasonal precipitation can be particularly strong in monsoon climates, such as in Asia. Drought impacts can be exacerbated by feedbacks between human activities and atmospheric circulation, such as deforestation, which leads to lower soil moisture retention.

Monitoring is the first step toward mitigating drought impacts, and several coordinated monitoring networks have been established (e.g., North America, the Mediterranean, sub-Saharan Africa, and Asia). However, in many cases mitigation and response actions are not taken until impacts are apparent. In contrast to other climate and weather-related hazards, drought is spatially pervasive, encompassing large parts of the landmass for months to years. Consequently, a variety of factors must be monitored, including water supply indicators (e.g., precipitation) and impact indicators (e.g., vegetation health, wildlife abundance, water quality). Some monitoring systems lack sufficient observations of key indicators, such as soil moisture or evapotranspiration, or a sufficient density of observations to gauge the spatial extent and severity of a drought to meet the needs of local decision makers. Remote sensing, a key component of drought and famine early warning networks, holds the promise of filling in monitoring data gaps. A key
challenge for improving the effectiveness of drought monitoring networks as an early warning tool is adequate and transparent provision of data and value-added information such as those provided in innovative research involving citizen scientists trained to report regional drought impacts.

Apart from improved monitoring, enhanced and reliable prediction frequently tops the request list of decision makers. At present, scientists have limited ability to predict drought, and the seasonal-to-annual drought prediction skill is primarily linked to the ENSO prediction skill. Even when the onset of a drought has occurred, the severity and duration of the drought are difficult to predict. Short-term forecasts, useful to agriculture and ecosystem management, may be derived from remotely sensed data, which may be include directly measured subsurface soil moisture, which is less dependent on precipitation, and thus can incorporate irrigation as a component of the surface water budget.

Seasonal-to-annual drought prediction is a function of initial hydrologic conditions and ocean-atmosphere forecast skill; thus, fruitful areas for research and model improvement include measurement or modeling of initial conditions, representation of air-sea interactions, translation of initial drought conditions into probabilities of impacts, and improved understanding of interactions between multiple ocean basins. On longer timescales, characteristics of Pacific and Atlantic tropical ocean-atmosphere circulation persist and may make possible the prediction of decade-length dry episodes and pluvials. Multiple ensembles of global circulation models are necessary for such predictions.

Social Construction of Drought

The conceptual understanding of drought has evolved in recent decades from a purely technical phenomenon based on hydrometeorological metrics to a largely social one. Political ecology research in geography approaches drought as a social construction, positing that drought risk is conditioned by a range of socioeconomic and political drivers. These include the gross domestic product, regional economy, patterns of water use and changing demand, land tenure regimes, government stability, and policy choices. Drought disasters are therefore subject to mitigation and societal adaptation. Vulnerability is a function of both exposure to drought risk and the capacity to adapt or make effective adjustments to socioeconomic systems in response to the impacts of drought. Geographer Diana Liverman, for example, demonstrated that access to agricultural technology and large landholdings reduced drought vulnerability relative to small, communal landholdings for farmers in Northern and Central Mexico. Vulnerability assessment methods may include participatory mapping (involving citizens and communities) and institutional capacity assessments. The institutional capacity of government agencies and civil society to anticipate drought impacts and respond via planning and mitigation strategies can reduce vulnerability to drought. Emerging participatory research models highlight the agency of citizens and institutions in anticipating and adapting to drought and involving key stakeholders in adaptive planning to reduce vulnerability.

Drought Planning

Drought planning can be an effective strategy to help a community, state, or country anticipate the challenges of drought and reduce the vulnerability of high-risk populations by establishing early warning mechanisms and coordinating responses of governmental levels and agencies. Australia’s 1992 national drought policy is considered emblematic for involving rural producers and balancing the competing demands of agriculture and the environment. In the United States, where there is a highly effective drought monitoring system at the national level, drought planning has been carried out largely at the state level. As of 2006, 37 U.S. states had developed state drought plans. Such plans generally frame drought severity levels and propose triggers of distinct actions to be taken at each level. Recent research has highlighted the contrast in drought and climate change preparedness in industrialized countries relative to developing countries; this gap needs to be addressed to reduce the high vulnerability of poor regions and protect human security.
Climate Change and Drought

What does climate change portend for the future of droughts worldwide? The Intergovernmental Panel on Climate Change suggests that the subtropics, low latitudes, and continental interiors are likely to experience more frequent and more severe drought through a combination of decreased annual precipitation and increased annual temperatures. Factors increasing the likelihood of drought vary regionally but include migration of storm tracks, increased evapotranspiration, and decreased soil moisture. Nonstationarity refers to an inherent dynamism or changing nature of climate associated with processes of human-induced climate change, meaning that the past climate does not serve as a reliable linear predictor of the future climate. Nonlinearities in the climate system (e.g., rapid continental ice sheet melt) can lead to abrupt climate changes or “climate surprises,” and nonlinear ecosystem responses can lead to dramatic changes in the landscape, evapotranspiration, and water quality. An example is rapid drought-induced forest mortality, with feedbacks that include substantial changes to insect pest life cycles.

The implications of climate change for future drought occurrence, prediction, and societal response are manifold. Traditional statistical methods of drought prediction may be confounded by nonstationarity, further challenging drought forecast efforts. Significant agricultural and ranching adaptations may be necessary to account for changing growing season lengths and temperatures. Water managers will need to seek additional sources of water, reuse treated effluent, or promote increased conservation as a way to cope with increased likelihood and severity of drought. Societies need to evolve globally a new adaptive sensibility that frames drought not as an inevitable disaster but as a hazard that can be managed to reduce vulnerability and build resilience.

Margaret Wilder and Gregg M. Garfin

See also Adaptation to Climate Change; Anthropogenic Climate Change; Biome: Desert; Desertification; Differential Vulnerabilities to Hazards; El Niño; Natural Hazards and Risk Analysis

Further Readings


Drugs, Geography of

Drugs comprise an array of chemical compounds that affect the human mind and body, and they occur in natural and synthetic, and licit and illicit forms. They are produced, exchanged, and consumed for numerous reasons, including for use in recreational and religious settings as well as medicinal purposes. Drugs and drug use are central to human health, welfare, and custom, and hold high cultural and economic value. At the same time, they can cause profound public health issues that often relate to drug abuse and dependence and also issues of criminalization. The geographies of drugs are therefore complex and contradictory. They operate at multiple scales that overlap and entwine, both linking and distancing people and places with and from each other, reaching through time as well as space. This entry begins with historical geographies of drugs in their various traditional, popular, commercial, medicinal, and illicit forms. It then looks at narcotics production under UN regulation; illicit drugs in the developing world; the growth in production of synthetic drugs; issues of drug addiction, related health problems, and treatment; and the use of spatial approaches in the analysis and policing of drugs.
Some drugs have very long traditions and uses associated with specific places. They include the ceremonial drinking of kava root by Pacific islanders, the ritualized ingestion of yagé or ayahuasca by shamans in South America, and the use of hallucinogenic peyote cactus by Midwestern American desert Indians and in the Native American Church. These traditional practices have tended to avoid strict legal control and criminalization but have instead suffered from disuse and marginalization. (Exceptions include the use of kat in Africa and the chewing of betel nut across Asia, both of which can be frowned on but are important in sociocultural and economic terms.)

Even traditional drug practices are regaining importance in their respective landscapes, however, and are significant for ecotourism and revenue creation, and for the reinvigoration of indigenous identities and cultural practices. Although such drug use does affect altered human states, it is far less problematic than modern drug abuse and rather is respected for reasons such as those cited above as well as for fostering local native knowledge and links back to more spiritual relationships with natural environments. Arguments for protecting pristine “Nature” and tropical rain forests, in particular, similarly promote renewed reverence for different peoples, places, and practices by claiming that they might hold the key to finding miracle cures for modern ills such as AIDS and cancer. Ethnobotanists therefore work with traditional healers and medicinal foragers in diverse, sometimes very small or isolated, communities as well as with large bioprospecting companies in seeking to convert folklore remedies into modern medicines (and large profits).

From a historical perspective, during the European “Age of Expansion,” for example, maritime explorers voyaged to Asia, Africa, and the Americas in pursuit of wealth from the spice trade and also in the hope of discovering plants with possible medicinal and culinary uses. The territorializing histories of colonialism that then followed subsequently featured valuable new drugs including quinine, caffeine, and tobacco. They have remained important cash crops for some local economies but still depend on transnational corporations, global markets, financial flows, and commodity chains. Such linkages also raise ethical questions as actors in affluent nations must decide whether or not to export products that might not be well distributed and are poorly regulated or costly in the less developed world, or whether to import cheap raw materials that sometimes rely on the exploitation of nonunionized and child labor or monopolistic markets. Alcohol is another drug with spatial manifestations shaped by political economy, history, and culture with attendant moral, legal, and health issues. Its geography connects (and disconnects) licensed venues, state legislation, dry Mormon states, Muslim nations, local temperance movements, and U.S. prohibition, for example, whereas wine regions, terroir, DOC appellation, and viticulture concern the more physical particulars of place, including soil, aspect, and climate.

Opiates used for the legal manufacture of analgesics likewise have their obverse in the drug heroin. India’s current position as the world’s primary producer of licit raw opium (exported mostly to the United States for processing) has a historical geography extending back at least to its Mughal rulers, the East India Trading Company, the growth of the British Empire, and Britain’s Opium Wars with China between 1839 and 1860. These narcotic drugs are also entwined in a complex geopolitics that has intensified in recent decades. The opium poppy is the oldest known medicinal plant and traces back millennia to around the Mediterranean and Asia Minor. With a Sanskrit name meaning “joy plant,” its main use has been to provide pain relief. Opium preparations were commonly taken until the 19th century, when alkaloids including codeine and morphine were first isolated. Then they and heroin, which was discovered in 1897 through a simple manufacturing process, became the main ingredient in remedies ranging from cough syrups to cure-alls. Addiction had previously been associated with, for example, Chinese opium smokers, British literary romanticism, or “morphinism” among American Civil War soldiers. It was seen as a more widespread problem in the 20th century as the moral panic over delinquency (embodying issues of youth and race), the criminalization of drugs, and black markets all boomed. As a result, several “wars on drugs” were announced...
DRUGS, GEOGRAPHY OF

in the 1970s and onward. They have been fought on various domestic and international fronts but are yet to be won.

Narcotics Production Since UN Regulation

The cultivation of poppies as well as coca and marijuana for legal narcotics production is licensed worldwide by the United Nations. Such regulation was first attempted at the International Opium Commission in Shanghai in 1909. Today, almost all nations are party to the United Nations’ key agreement of the 1961 Single Convention on Narcotic Drugs (as amended by the 1972 Protocol) as well as the 1971 Convention on Psychotropic Substances and the 1988 Convention against Illegal Traffic in Narcotic Drugs and Psychotropic Substances, which have been agreed on to control new problems such as the development of synthetic drugs, global drug cartels, and international money laundering. The main regulatory body is the International Narcotics Control Board, which works with the UN Office of Drugs and Crime. These conventions are mostly prohibitive. However, the United Nations does not have the power to enforce implementation or punish noncompliance, which remains under the domestic jurisdiction of individual parties. Most important are the multilateral agreements, which determine the areas to be cultivated and the quantities of licit narcotics to be produced annually for each nation by comparing the global stocks with estimated demand and by balancing imports and exports.

Traditional opium production involves scraping latex from the incised green poppy capsules by hand, and this labor-intensive agriculture has supported millions of peasant farming families from the Middle East to the Indian subcontinent. Its legitimacy has waned though, and the modern “poppy straw process” is now most common and encouraged under UN licensing. The process is highly mechanized and relies on industrial chemical extraction, and thus world production of licit narcotics is now led by nations including Australia, France, Hungary, and Spain. However, the UN ensures that two major traditional producers are still accommodated. India is therefore uniquely licensed to continue exporting opium, whereas Turkey, which was banned from production in 1972, underwent modernization with relicensing 2 years later and now leads global production with its output of “concentrate of poppy straw” (Table 1). Also, the United States is the world’s biggest importer and consumer of licit narcotics. In 1980, it passed the “80/20 rule,” requiring that at least 80% of its imported raw narcotic materials come from India and Turkey. This legislation concerns both opium and poppy straw but forms of the latter that are most usually rich in morphine. However, it overlooks a new form of poppy straw rich in the alkaloid thebaine (rather than morphine). The official reporting of thebaine production for the United Nations began in 1999. Reports reveal rapid increases, which they explain in terms of the growing international demand for particular pharmaceutical drugs such as oxycodone and buprenorphine (painkillers so popular that the former now has a name, “hillbilly heroin,” in the United States). Australia has led global production, accounting in 2005 and 2006 for 70% and 80%, respectively, of the total 70 tons and 72 tons of thebaine equivalent produced those years.

In regulating narcotics, the United Nations tries to guarantee the world’s licit supplies but also aims to address the problems of the illegal drugs trade. Licensing a modernized poppy industry has lessened the crop diversion, for example, which was once rife in India and Turkey. Government legislation, official inspections, and policing at various levels are also central to its security. The Australian industry, as a relative newcomer, even confines its poppy cultivation to the island state of Tasmania and until very recently was dominated by just two international companies following its establishment in the 1970s. Narcotics production is most profitable for licensed farmers across the world complying with regulations. However, the illicit cultivation, manufacture, and trafficking of illegal drugs that are part of the international organization of crime (and sometimes terrorism) can resist such control. Such farmers have links especially to fragile states struggling with political instability, armed conflict and dire poverty.

Lure of Drugs for the Developing World

In the 1950s, communist prohibition pushed poppy cultivation from China to southeast Asia
to Central Asia, where the infamous “Golden Triangle” (Burma, Laos, Thailand) and “Golden Crescent” (Afghanistan, Iran, Pakistan) are found. With the Cold War and increased militarization in these regions, opium and heroin manufacture were not suppressed but simply shifted to become involved in the political conflicts and related financing of local war efforts. In Burma (now Myanmar), both the ruling junta and the autonomous ethnic groups backed by the United Wa State Army rely on and negotiate ceasefires around their illegal drug production and trade. The U.S. provision of foreign aid to such nations is most often tied to the eradication of illicit crops. However, the CIA’s backing of resistance groups that include the Hmong in Laos, the Nationalist Chinese or Guomindang in Northern Burma, and the Islamic mujahideen in Afghanistan has implicated them in the continued and increasing production of opium and heroin in these two regions. While poppy cultivation had been forbidden under the Shah in Iran since the 1950s, it is now, like Pakistan, a place instead for processing and trafficking. Likewise, in Afghanistan, following the Soviet withdrawal, the Taliban reduced opium production there from 3,300 tons in 2000 to a miniscule 185 tons in 2001 (but was still trading from stockpiles).

Afghanistan’s return and rapid rise to lead the world in opium production has reached record-breaking levels each year since 2006. This illicit economy involves more than half a million households with a total farm gate value there of US$1 billion, and its opiate exports to neighboring countries in 2007 have been put at US$4 billion, constituting an estimated 33% of Afghanistan’s GDP. International commentators have therefore investigated the prospect of Afghanistan becoming a licensed producer. However, the suggestion is hampered by problems of poverty, food scarcity, political conflict, and the inadequate progress made with crop substitution programs and infrastructure development, including the repair of irrigation systems completely destroyed by war. More than half the illicit production in this country now occurs in the one southern province of Helmand, where the Taliban has a strong presence. With the movement southward of production to there, trafficking has also shifted from the old Silk Route to the Balkan route. Therefore, while heroin seizures are increasing in Russia and other former Soviet states, so are institutionalized crime and corruption, heroin addiction, and HIV infection rates.

Patterns of drug production and trafficking (and, less so, consumption) change quickly. Coca cultivation, associated with South America, is appearing on islands in the Caribbean and Oceania, for example. Diversification of drug economies is occurring worldwide, and opiates are now being grown in addition to coca and marijuana in Colombia, for example. There the drug lords and traffickers (cocaleros) have had alliances with the revolutionary army, FARC, since the early 1980s,

<table>
<thead>
<tr>
<th></th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
<th>2008</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia</td>
<td>151</td>
<td>96</td>
<td>130</td>
<td>70</td>
<td>68</td>
<td>53</td>
</tr>
<tr>
<td>France</td>
<td>68</td>
<td>101</td>
<td>96</td>
<td>56</td>
<td>25</td>
<td>49</td>
</tr>
<tr>
<td>Hungary</td>
<td>9</td>
<td>30</td>
<td>15</td>
<td>17</td>
<td>15</td>
<td>49</td>
</tr>
<tr>
<td>India</td>
<td>57</td>
<td>92</td>
<td>37</td>
<td>38</td>
<td>30</td>
<td>24</td>
</tr>
<tr>
<td>Spain</td>
<td>44</td>
<td>55</td>
<td>36</td>
<td>17</td>
<td>70</td>
<td>85</td>
</tr>
<tr>
<td>Turkey</td>
<td>145</td>
<td>60</td>
<td>64</td>
<td>106</td>
<td>30</td>
<td>70</td>
</tr>
<tr>
<td>Others</td>
<td>13</td>
<td>13</td>
<td>13</td>
<td>12</td>
<td>14</td>
<td>22</td>
</tr>
<tr>
<td>Total</td>
<td>487</td>
<td>447</td>
<td>391</td>
<td>316</td>
<td>252</td>
<td>352</td>
</tr>
</tbody>
</table>

Table 1: Global production of licit raw narcotics rich in morphine in the years 2003 to 2008. The world’s major producers of morphine-rich narcotics include several European countries as well as India and Australia.

wielding great power through political corruption and “narcoterrorism.” However, dismantling of the Medellín and Cali cartels through the 1990s revealed a capacity for dynamic reorganization. Coca paste continues to be produced, almost all for export as cocaine, but then it was also being marketed as “base” or crack, triggering a new epidemic in the United States that affected black urban neighborhoods already in decline. The method of its transport also changed from air to maritime routes. And as coca production increases in Colombia, Bolivia, and Peru, trafficking to Europe has incorporated more West African transit points, which once focused only on the local production of marijuana and hashish. Mexico has similarly become a major processing and transit point for different drugs destined for the United States and elsewhere. Dynamic drug geographies are also evident, with the decriminalization of marijuana in Amsterdam spreading to some Australian and U.S. states too. Indeed, the focal point that California once played in the psychedelic era of the 1960s and 1970s has lessened, and the geography of LSD or acid has since been reconfigured along with many other, newer synthetic drugs.

**Synthetic Drugs**

The manufacture of synthetic drugs is also meeting new demands in the illicit economy.
Chemically manufactured drugs such as amphetamines or “speed” have had links to British and North American rebellious youth cultures of rock ‘n’ roll and motorcycles for decades. They once relied on the diversion of pharmaceuticals such as benzedrine, but nowadays a new industry that manufactures an array of illegal synthetic drugs is meeting the increasing demand here. Ecstasy pills (“E”) are easily produced for illicit markets in bulk, for example, and are then sold with international brand names known to consumers. They are often used en masse at dance parties and rave scenes and has been associated most with the British independent music coming out of Manchester in the 1990s as well as with select, international tourist locations for young people, ranging from Ibiza to Goa. Other designer drugs that are now appearing include “fantasy” and “GBH.” Likewise, the production of ATS or “amphetamine-type stimulants” such as methamphetamine in Burma is fuelling a new drug consumption culture in Thailand. The use of methamphetamine, known worldwide as “ice” or as yaa baa (meaning “madness”) in southeast Asia, is associated with students and workers as much as with partygoers and nightclub but still threatens serious psychological, generational, and socioeconomic damage. Related factors in controlling the problem’s spread are the small size, mobility, and number of drug factories that can now easily and quickly change location so as to evade detection by the authorities, with minimal disruption of their illegal operation.

**Addiction, Health Problems, and Treatment**

Like the problems that they are designed to address, public health strategies around drug use exhibit distinct and dynamic spatial patterns. Harm reduction policies constitute an alternative to the usual efforts at control through criminalization and policing of drugs (their supply, in particular). Radical versions have included free prescription drugs, public injecting rooms, and acceptance of “heroin parks” variously in the Netherlands, Switzerland, and the United Kingdom. But more generally, they involve education on safe drug use practices, provision of needles and syringes, peer support, counselling, and rehabilitation and reduction or maintenance of drug use through substitution programs (especially using the synthetics methadone in the United States and Australia and buprenorphine in France to treat addiction to opiates). They have been perhaps most successful in the United Kingdom, Canada, and Australia in significantly reducing the incidence of HIV/AIDS and hepatitis B and C infections as well as other drug-related morbidity and mortality rates. The uptake of such policy and practices varies greatly between national and subnational jurisdictions and even at the state and county levels. Therefore, in examining the epidemiology of addiction and the problems of drug epidemics, HIV/AIDS, and overdose fatalities, as well as possible solutions, geographers have looked at other factors. In addition to the role of local drug economies and subcultural practices, they include the location, type, and accessibility of service delivery; the influence of the built environment; housing and tenure type; and community factors such as race, ethnicity, income, and transience. Still predominant though is a belief in the broader approaches and their capacity to influence supply-side controls through legislation and policing in conjunction with anti-drug campaigns to lessen demand.

**Spatial Approaches to Analysis and Policing**

Early-20th-century geographies of social and spatial disadvantage, including social area analysis and urban ecology, sometimes included drug use or delinquency as a variable. The geography of drugs has also developed with the use of environmental approaches to crime. Using “routine activity” and “social disorganization” theories; mapping the spatial distribution of crimes; analyzing hotspots; constructing geographical offender profiles, distribution networks, and mobility patterns; and applying traditional tools such as “distance decay,” “spatial diffusion,” and “journey-to-work” models have all helped describe where and how different drug markets operate. But the geographical understandings of crime and policing that are most pronounced in policy and practice are situational crime prevention and crime prevention through environmental design. They are now being similarly applied to the problems of drugs. Initiatives in the United States, for example, involving increasing the
Dunes are depositional landforms, comprising accumulations of sediment (sand, silt, or clay), typically deposited by wind (eolian) action. They may be defined according to the type of sediment they comprise (sand dunes, clay dunes), their morphology (linear dunes, barchan dunes, star dunes, parabolic dunes, transverse dunes), their position relative to other geomorphic features in the landscape (barrier dunes, lunette dunes, headland bypass dunes, climbing-falling dunes), the time of their formation (contemporary dunes, paleo- or relict dunes), or even according to their genesis (natural dunes, artificial dunes).

Dunes vary in extent and, typically, reach heights of between 5 m (meters) and 30 m. Mega-dunes reach 100 m or, in extreme cases, 400 m or more in height. Dunes occur typically, but not exclusively, in hot deserts and coastal environments. The sediment comprising a dune is usually quartz sand, with a carbonate component in coastal environments and a gypsum component in certain arid environments. The formation of sand dunes requires a plentiful supply of sand and a consistent wind direction. An obstacle, commonly vegetation, may retard the flow of air near the surface, initiating deposition. Once a

---

Stewart Williams

See also Carcinogens; Disease, Geography of; Health and Health Care, Geography of; HIV/AIDS, Geography of; Indigenous Environmental Knowledge; Law, Geography of; Medical Geography

---

Further Readings


Dune is established, it will continue to grow, provided that the upwind sand supply is maintained. Its ultimate height and form are a product of factors described below.

Dunes may have a complex internal structure, particularly where the wind direction is highly variable. Dune morphology (shape or form), internal structure (bedding), and composition (the nature of the sediment comprising the dune, and its inherent physical and geochemical properties) reveal evidence regarding the formation and growth of the dune and regarding the source of its sediment. Ground-penetrating radar may reveal the internal structure of the dune. Radiometric dating techniques such as optically stimulated luminescence (OSL) facilitate the determination of the time when the dune sediment was last exposed to sunlight. Effectively, this is the last time the dune was active.

The morphology of dunes underpins the most common means of classification. The amount of sand available, variability of wind direction, wind strength, presence or absence of vegetation cover, and nature of the terrain interact to determine the type of dune that will form. Much of our understanding of dune formation and dune-building processes is derived from the classic work of R. A. Bagnold in the 1940s. When wind blows over a surface, frictional resistance between moving air and the surface results in a complex pattern of turbulence, vortices, and eddies. Dune building is the result of an intricate and dynamic interplay between wind, sediment, and obstacles or irregularities on the surface.

Barchans are the most common dune type, and so their formation is described here as an example of dune processes. Barchans form where there is a limited sand supply and sparse vegetation cover. Barchans are asymmetrical, with a gently sloping (~5° to ~10°) windward (stoss) face and a steeper (~30° to 35°, which is the angle of repose of dry sand), concave lee or slip face. The horns of these crescent-shaped dunes point downwind. Under unidirectional wind conditions, sand grains advance up the windward face of the dune through a process of short hops (saltation). As the top of the lee face oversteepens through the accumulation of more and more sand, it eventually fails and sand moves down the face till the angle of repose is reestablished. The process repeats itself as long as there is a supply of wind-blow sand off the windward face, onto the top.

Figure 1  Dune cross-section. Note successive slip faces, the product of sand migration up and over the windward slope.

of the slip face. The result is twofold; clearly defined foreset beds occur on the slip face, and the dune advances downwind. Barchans may migrate downwind over long distances. Figure 1 shows a dune cross-section.

Transverse dunes are formed where the sand supply is more abundant, and individual barchans coalesce. Where the crests are sinuous, these dunes are referred to as barchanoid ridges. Crests are at right angles to the prevailing wind. Linear or longitudinal dunes are, as their name implies, long and straight. They often form under a bidirectional wind regime. Dune crests run parallel to one another and to the wind. These knife-edge crests give rise to the alternative name seif (Arabic meaning “sword”) dune.

Star dunes form when multidirectional winds are prevalent. Their radiating arms are a characteristic feature.

Parabolic dunes are U-shaped, with the arms pointing upwind. They are common in coastal environments where the wind maintains a constant direction off the ocean. A complex array of parabolic dunes, with multiple crests and slip faces, results. Blowouts often occur on the upwind side of such dunes.

Lunette dunes have a half-moon (crescent) shape and occur on the downwind side of playas (depressions periodically occupied by standing water to form a temporary lake). Sand or clay is deflated from the playa surface or shore and accumulates above the shoreline.

Dome, star, and reversing dunes are derivates that occur when the wind direction is either multidirectional or highly variable.

Figure 2 shows the various dune types. The world’s hot, dry deserts, in particular the Sahara, Arabian, Atacama, Namib, North American, Great Australian, and Gobi deserts, are the repository for a variety of dunes on a vast scale. However paleodunes, often stabilized by vegetation, are found in both tropical and temperate climatic regions. They are presumably an artifact of past climate change. The reverse holds true: Areas of sub-Saharan Africa that comprised scrub or grassland half a century ago are now covered by mobile sand dunes.

Sand seas or ergs are found in the Sahara, Namib, Arabian, and Gobi deserts. They comprise vast areas of mobile sand, often covering tens of thousands of square kilometers. A variety of dune types occur, though some sand seas are dominated by a particular type, for example the linear dunes of the Namib sand sea. The mobility of the dunes produces a constantly shifting, hostile landscape.
Coastal dunes are common along sandy coasts. They are intimately connected to beaches, which serve as a source and sink for sediment, and beach processes. Wave action deposits sand onshore, from where it may be reworked by the wind, thereby nourishing the coastal dunes.

The failure to recognize the dynamic nature of dunes can be disastrous. Settlements, infrastructure, and productive land may be smothered by migrating dunes. The stabilization of dunes can have unintended consequences. For example, alien Australian species introduced to South Africa to control shifting dunes now threaten the plant diversity of the unique Cape *fynbos* (fine bush) floristic kingdom. On the positive side, dunes may serve as a source of building or glass sand. Sedimentological evidence from paleodunes or fossilized dunes can be used to reconstruct past environmental conditions, thus abetting an understanding of climate change.

Peter J. Holmes

See also Biome: Desert; Coastal Erosion and Deposition; Desertification; Environmental Management: Drylands; Geomorphology; Global Environmental Change; Landforms; Wind Erosion

Further Readings


The term *dynamic display* is applied to displays capable of *change*, in contrast to static displays, such as paper maps and graphics. Changes in a dynamic display may occur in an automatic manner or in response to actions of the user. The term *interactive display* is applied to dynamic displays that change in response to a user’s actions. Dynamic displays capable of changing automatically are often called *animated displays*. An animated display may at the same time be interactive. At the very least, the user may interactively stop, resume, and restart the play, but more advanced interaction possibilities may also be provided.

Computer animation is often used in a straightforward way to display dynamic phenomena. In particular, dynamic phenomena related to geographic space are most often visualized by means of animated maps. Many examples of such maps may be found on the Web. Thus, animated maps depicting various topics of American history are available at AnimatedAtlas.com. The Web site of the National Atlas of the United States contains animated maps portraying the spread of invasive species during a range of years and the progression of vegetation growth during a year.

Another frequently encountered class of animated displays is fly-through displays, which present various objects or geographical areas as they appear from different viewpoints. Typically, fly-through displays allow very little or no interactivity. Animation may also be used to present large and/or complex content by portions. For example, different themes may alternate in an animated map of a geographic area: terrain, rivers and lakes, landmarks, vegetation, administrative boundaries, roads, and so on.

Interactive displays and, in particular, maps and virtual globes are nowadays frequently encountered on the Web. Well-known among these are Google Maps and Google Earth, online mapping and route planning services such as Map24, and maps displaying election results (e.g., those at NYTimes.com). Typically, interactive maps and globes (as well as other types of display representing large amounts of data and complex data) provide facilities for view navigation. The most basic navigation tools are zooming and panning. Perspective views of three-dimensional objects or spaces provide tools for controlling the user’s viewpoint in the scene. Visualization researchers suggest “intelligent” navigation techniques that adjust the view to the user’s focus of interest so that the visual prominence (in particular, size of graphical elements) and level of shown detail vary depending on the relevance of the information in different parts of the display for the user.

Another typical feature of interactive displays is providing additional information about objects or places in response to clicking on graphical elements representing these objects or places. The information may be shown in a pop-up window that appears near the cursor position, in a dedicated area of the display or in an adjacent subwindow.

Thematic maps, such as tourist maps and maps of election results, often support more advanced interaction, allowing the user to alter the information content of the display or the representation of the information. Examples of such interaction include the following:

- Selecting information layers (e.g., accommodation, museums, restaurants, bar/nightlife) on a tourist map or toggling the visibility of the layers
- Selecting time moments or intervals in displays of time-related information
- Changing the spatial scale (e.g., states or counties)
- Changing the theme (e.g., president elections or governor elections, absolute values or differences in comparison with the previous time)
- Choosing the visualization method (e.g., choropleth map or proportional symbols)

Interactive displays used for specialized services may be supplied with specific interaction facilities, such as selection of source, destination, and intermediate places in route planning.

Interactive displays play a very important role in the exploration and analysis of data and information. Typically, the size, dimensionality, and/or other complexities of data that need to be analyzed preclude the possibility of simultaneous

(Continued on page 808)
Figure 1 A map (A) is dynamically linked to a time graph display (B) by simultaneous highlighting of corresponding graphical elements.

Source: Gennady Andrienko.

Note: The user has selected several states on the map; the states are designated in yellow. Simultaneously, the lines on the time graph representing the temporal dynamics of the property crime rates in these states are shown in black.
Figure 2  Dynamic filtering
Source: Gennady Andrienko.
Notes: The map (A) shows the spatial positions of earthquake epicenters; the perspective view (B), where the vertical dimension represents time, shows the positions of the earthquakes in space and time. Both displays show the same subset of the data selected by means of three different filters: (1) spatial window, which has been drawn by the user within the map display; (2) attribute filter (C), which selects the earthquakes with the magnitudes 4 or more; and (3) temporal filter (D), which selects the earthquakes that occurred in the period from the beginning of 1995 till the end of 1999. All three filters may be interactively changed by the user; in response, the displays will immediately change their content to satisfy the new filter conditions.

Figure 3  The process of interactive change of the color scale in a choropleth map
Source: Gennady Andrienko.
Notes: On the left, the percentages of female among unemployed people in the districts of Portugal are represented on a map by means of a diverging, or double-ended color scale; shades of blue represent values below 50% and shades of brown are used for values greater than 50%. Moving the slider along the color bar on the right of the map changes the reference value and, hence, modifies the color scale. The middle and right screenshots correspond to the reference values of 60.25% and 67.31%, respectively.
Figure 4  Interactive data transformation applied to an animated map of the population dynamics in Germany. 
Source: Gennady Andrienko.

Notes: Each row contains screenshots of the states of the animated map corresponding to three consecutive years. Upper row: The map represents the absolute population numbers. Middle row: The data have been transformed from the absolute values to the differences between two consecutive years. The blue circles represent negative differences. Lower row: The same data have been transformed to ratios between consecutive years. The blue circles represent ratios below 1, the red circles are ratios greater than 1.
representation of all data items, dimensions, structural parts, relationships, and so on in a single display. Hence, the analyst has to understand the whole by looking at subsets, components, projections, and aspects of the data. Interactive displays enable a dialogue between the analyst and the data. A fundamental interactive operation required for exploration and analysis is focusing, that is, selection of subsets, projections, and so on for viewing. Since any view can only convey partial information, the analyst needs multiple views, which need to be linked so that the pieces of information contained in them could be fit together. Thus, multiple displays simultaneously present on the screen may be linked through identical marking of corresponding graphical elements, for example, with the same color. The best-known technique, often referred to as “brushing,” is simultaneous highlighting of graphical elements corresponding to objects selected by the user in one of the displays (Figure 1).

Another well-known technique that supports both focusing and linking is dynamic filtering (also known as conditioning): All displays show only the data satisfying user-specified conditions and dynamically update their content when the user interactively changes the conditions. Data may be filtered according to the values of one or more attributes as well as according to their spatial and/or temporal references (Figure 2).

Apart from focusing and linking, the role of interactive displays is to help the analyst discover unobvious patterns in data. Two major groups of interaction techniques intended for this purpose are display manipulation techniques and data manipulation techniques. Display manipulation is interactive modification of the method of encoding elements of the data by graphical elements and their properties. For example, Figure 3 demonstrates the process of interactive change of color scale in a choropleth map. This operation allows the analyst to explore more comprehensively the spatial distribution of the attribute values.

Data manipulation is interactive transformation of the data elements represented in a display. For example, Figure 4 demonstrates interactive transformation of time-referenced data. Initially, the animated map represents the original data values corresponding to the geographic areas and consecutive time moments. These values can be interactively converted into differences or ratios with respect to the previous time moments. The resulting map exposes previously unnoticeable spatial patterns.

Gennady Andrienko and Natalia Andrienko

See also CAD Systems; Cartography; Kwan, Mei-Po; Map Animation; Map Visualization

Further Readings


Carville Earle was a historical geographer whose primary goal was to answer vexing historical questions of the American past using the tools and methods of geography. Earle’s research spanned from colonial Jamestown to the 21st-century United States. He often referred to himself as a geographical historian practicing geographical history. He was a pioneer in using geographic information systems in the cause of answering problems of the past with the Historical United States County Boundary Files, which allowed scholars to map and analyze historical data at the county level. Earle also served as editor of the *Annals of the Association of American Geographers* from 1994 to 1996.

Earle’s major contributions to historical geography emerged from his insistence that historical geographers were social scientists with a focus on using a geographic or environmental approach to studying problems of the American past. Historical geography could provide new insights and answers to historical questions. Earle and Donald Meinig sparked debate among historical geographers on what direction and guidance historical geography needed. Earle critiqued what he called Meinig’s continuous imperialistic expansionist interpretation of the American past. He believed that constant upheaval and change were caused by the ravenous appetite and cyclical nature of capitalism. He saw the American past as a process of economic stops and starts and of crisis and recovery that were manifested and interpreted on the settled landscape. His book *The American Way* highlighted the cyclical nature of the American past based on Kondratieff long waves and a 50-year cycle of government policies in response to the cyclical nature and expanding influence of the American state in response to crisis in capitalism. Methodologically, Earle’s greatest influence will be his advocacy of spatial and empirical methods. Theoretically, he contributed to the understanding of alternating policy regimes of the American state and how they are manifested spatially on the landscape.

His major publications included *The Evolution of a Tidewater Settlement System: All Hallow’s Parish Maryland 1650–1783* and *Geographical Inquiry and American Historical Problems*, which included previously published seminal journal articles and new research. *Geographical Inquiry and American Historical Problems* included the influential pieces “Environment, Disease, and Mortality in Early Virginia,” “Staple Crops and Urban Development in the Eighteenth-Century South” with Ronald Hoffman, “A Staple Interpretation of Slavery and Free Labor” and “Socialism in America: A Geographical Interpretation of Its Failure” with Sari Bennett. His final work was
Earthquakes involve the travel of elastic, or seismic, waves of energy through the Earth’s rigid crust. They result from the sudden liberation of accumulated stress and strain along faults, which are planar discontinuities in the crust. Around the world, more than 3,000 seismic events occur each year, but few of them are damaging or lethal. However, when major earthquakes occur in highly populated areas, they can cause major destruction and death tolls in the thousands.

**Distribution of Seismicity**

The global distribution of seismicity is conditioned by the pattern of tectonic plates, which is a function of patterns of convection in the Earth’s mantle and variations in the planet’s spin rate. Most earthquakes are concentrated at the plate margins and in particular around the subduction zones, where crustal material is subsumed by overriding plates. Transcurrent plate margins also generate significant seismicity, as occurs along the west coast of North America, where the Pacific Plate is moving north against the continental plates to its east. Collisional plate margins generate seismicity through their impact on orogeny, or mountain building. In such regions, seismicity is usually associated with landslides. Finally, the regional geometry of stress creates some “hot spots” of intraplate seismicity, which are not associated with the plate margins. Particular problems are experienced with “seismic gaps”—sections of plate margins, or other seismogenic areas, where strain has built up and not been released in a major earthquake for decades or centuries.

Owing to the world distribution of the human population, the pattern of human effects of earthquakes does not coincide exactly with the global pattern of seismicity. Major earthquakes occur in unpopulated areas (e.g., in Alaska and Siberia) with limited consequences for human settlement. Much less severe earthquakes can have terrible human consequences if they occur in areas of dense settlement.

**Earthquake Mechanisms**

It is a common misconception that earthquakes can be characterized adequately by a single measure. This is usually given as magnitude or intensity, but in reality more information is needed. Magnitude is a surrogate measure of the energy liberated in the shaking, known as strong motion. Modern scales owe much to the pioneering work of the Californian seismologist Charles F. Richter. However, although his name is widely used in the mass media, the Richter scale has been abandoned by seismologists as it tends to be inaccurate at high magnitudes. Instead, a combination of measures is used that characterizes the magnitude of different seismic waves, particularly body waves, which travel through crustal material, and surface waves, which travel along surfaces and discontinuities in rock formations. The six kinds of seismic waves (i.e., P waves, S waves, body waves, surface waves, Rayleigh waves, and Love waves) involve different forms of motion. For example, the first waves to arrive (P waves) at a point on the Earth’s surface are compression-extension movements of a longitudinal kind, and these are followed by the undulatory motion of S waves.

In terms of surface effects and damage, two other measures are vitally important. The first is hypocentral depth, the distance below the surface at which earthquake shaking begins. In most geophysical models of earthquake generation, intensive microfissuring coalesces to create sudden variations in stress within a block of rigid crustal
material. As stress and strain are liberated, shaking spreads out dynamically from the small area in which the earthquake hypocenter, or focus, is born. Hypocenters that are deeper than 50 to 80 km (kilometers) seldom generate damaging earthquakes, and so the lethal seismic events tend to be those that have shallow foci, perhaps 5 to 20 km below the surface. In the Tonga Trench, in the South Pacific, major earthquakes occur as deep as 700 km, but by the time the seismic waves reach the surface they have dispersed their energy and are harmless. Conversely, the earthquake that occurred in Bam, Southern Iran, in December 2003 had a hypocenter only 10 km below the surface, and it killed 26,271 people.

A further measure of importance is bracketed duration, the length of time that strong motion lasts. Whatever its magnitude, an earthquake of 10 to 15 s (seconds) is likely to do less damage than one that lasts a full minute and is therefore fully able to propagate cracks in buildings. In Mexico City in 1985, shaking lasted intermittently for 5 min. (minutes), and the damage was considerable. However, many engineers regard the most useful descriptor of an earthquake’s power to be seismic acceleration. This is measured or calculated as a proportion of the universal gravitational constant, \( g = 9.81 \text{ cm/s}^2 \) (centimeters per square second). In major earthquakes, acceleration can exceed 1.0g, although most antiseismic construction is predicated on values of 40% or 60% of g. Acceleration depends on both the crustal adjustment forces liberated during the earthquake and the type of rocks through which the seismic waves travel. It can be particularly strong in soft sediments, such as sands and coastal infills.

In addition, the frequencies of seismic waves are an important means of characterizing an earthquake. Particular frequencies interact with buildings, bridges, and vehicles in especially destructive ways. Each building has a fundamental period (the inverse of frequency), as if it were a bell that resonates at a particular pitch when struck. Given the additive properties of waves, if the earthquake’s mix of frequencies matches the fundamental period of a building, shaking effects will be amplified. Moreover, if the Japanese Shinkansen high-speed train is struck laterally by seismic waves with a dominant frequency of 56 Hertz it may be rolled off the tracks, as happened in the 2004 Chuetsu earthquake in Niigata Prefecture.

Finally, seismic intensity is a measure of the felt effects and damage caused by earthquakes. As these depend on perception (at low intensities) and the resistance of buildings to damage (at high intensities), intensity values tend not to be correlated very strictly with magnitude or other physical measures. Hence, a relatively weak earthquake can cause very serious damage if building stock is not resistant to seismic shaking. In Bam in 2003, an earthquake of magnitude 6.6 caused extensive high-intensity damage in weak-walled masonry and adobe buildings, which were not sufficiently resistant. Modern intensity measures stem from the work of the Italian seismologist Giuseppe Mercalli (1850–1914). In the Americas, the Modified Mercalli (MM) scale is used, while in Europe the Mercalli-Cancani-Sieberg and Medvedev-Sponheur-Karnik scales are preferred. The difference is a function of the differing building techniques that are used in the two continents.

### Measuring Earthquakes

The process of seismic measurement is less precise than it would seem to be. Despite advances in laserometry and digital measurement, seismometers (continuously recording devices) and accelerometers (which are triggered by strong motion) must still achieve a difficult balance between sensitivity and robustness, just as they had to when the first modern measuring apparatus was invented more than 100 years ago. An epicenter, the location on Earth’s surface directly above the original source of shaking, the hypocenter or focus, is in reality an ellipse, not a point, as it can only be determined approximately. It is not necessarily the location of maximum damage, as this will also depend on the configuration of topography and rock formations, across which the waves are transmitted, and the patterns of human settlement and land use.

### Physical Effects of Earthquakes

Earthquakes can be major sources of mass movement or landsliding. Large seismic events can generate as many as 15,000 landslides, largely confined to the area of intense shaking, perhaps
MM intensity VI and above. Some landslides can be extremely large; however, few types of mass movement are largely or exclusively set in motion by earthquakes. Liquefaction, block glides, and lateral spreads, involving plastic-fluid movements at very low angles, are characteristic of seismic effects. Most other movements, from mudflows to slumps, may be activated or reactivated by earthquakes but can equally occur in the absence of seismicity, albeit seldom in such copious quantities.

One common effect of major earthquakes in mountainous areas is to cause landslide dams on rivers. The sequence of six earthquakes that occurred in Calabria, Southern Italy, in 1783–1785 created 211 landslide-dammed lakes. However, when landslide debris invades a river channel and blocks it, the result is likely to be not a permanent lake but a sudden outburst flood. Permeable, unconsolidated debris, increasing pore pressures, and the weight of water as it accumulates all conspire to cause breaching within 2 weeks of the blockage. Such an event occurred on the flanks of Mount Huascarán in Peru after the May 1970 earthquake, and only prompt action by the Chinese authorities stopped it from happening after the Sichuan earthquake of May 2008. Finally, earthquakes can reroute rivers. This has frequently occurred in the deltaic sediments of Bangladesh, which has 254 rivers and occasional high seismicity derived from the nearby Himalayan orogeny.

Tsunamis, which are sometimes called seismic sea waves, may be generated when earthquakes occur at the coast or beneath the seafloor. Generally, a shallow-focus event of magnitude of at least 7 must be involved—that is, a powerful earthquake. Not all undersea events cause tsunamis as there must be an abrupt movement of the whole body of seawater, usually in a vertical
direction, perhaps by sudden displacement of the seafloor and hence the column of water above it. Energy moves out from the source with an elliptical motion that propagates water waves of low amplitude and long wavelength. When these reach the shallow water of the coast, the principle of energy conservation dictates that friction on the seabed and the associated velocity reduction will cause the wave to rear up (in about one third of tsunamis the trough arrives at the coast before the first crest, causing the sea to retreat for several minutes). Waves can be refracted around islands and funneled by inlets such that a coastal amplitude of 3 or 4 m (meters) may be increased by an order of magnitude in specific places. A tsunami can have a devastating effect on low-lying and densely populated coasts. Given the huge distances and long travel times in open oceans, the opportunities for warning and evacuation are substantial, but the greatest risks occur with near-field tsunamis, those that are generated locally and in which the time between inception and arrival at the coast may be limited to a few tens of minutes, not several hours.

The earthquake of December 26, 2004, in the subduction trench off the coast of Indonesia caused a devastating tsunami that greatly affected 12 countries bordering on the Indian Ocean basin. Wave heights at run-up exceeded 9 m in some areas and over-wash obliterated large sections of coast, for example, in Banda Aceh, Indonesia and Sri Lanka. The death toll was 292,000, largely through drowning.

Human Effects of Earthquakes

Each year between 7 and 20 earthquakes cause deaths and injuries. Tens more cause damage to buildings and structures. The worst effects occur where high concentrations of population coincide with powerful seismicity, for example, in China, Indonesia, Iran, Afghanistan, and Japan. The January 2010 earthquake in Haiti killed more than 200,000 people.

Specialists in earthquake injury epidemiology have postulated a ratio of one death per three significant injuries, but whereas there are earthquakes that produce this sort of pattern, data
tend to be highly variable from one event to another. This reflects the complexity of various factors. The first of these is the seismic resistance of building stock, which depends on construction type and on state of maintenance. Generally, well-maintained frame structures built according to antiseismic criteria perform well, whereas dilapidated rubble masonry dwellings that incorporate no seismic protection tend to collapse easily. Despite this, the ratio of deaths to collapsed buildings appears to be low, perhaps in the range of 12 to 33 per 100 buildings that totally collapse. If this is so—and research is inconclusive—it may reflect both building occupancy rates and the impact of self-protective behavior.

Another factor whose impact on death rates is poorly understood is search and rescue. Although healthy people who have not been injured can survive trapped under rubble for up to 2 wks., the vast majority of people who are rescued from collapsed buildings are brought out within a matter of hours after the earthquake strikes. Injury can substantially reduce survival times, especially if it involves loss of blood, vital functions, or body heat. One common medical effect of entrapment is the crush syndrome, in which the rupture of cells liberates potassium and lipids into the bloodstream and causes kidney failure. People rescued alive with crush injuries may need dialysis.

Field experience suggests that serious injury in earthquakes can take many forms. Whereas paraplegia and the need for in situ amputation are rare, multiple fractures, cranial trauma, and compression of the thorax are common. In many cases, minor injuries are associated with flight and result from falling down stairs (broken or sprained limbs) or walking on shards of glass (lacerations).

Currently, the greatest risk of casualties from earthquakes exists in several of the world’s “megacities,” notably Tehran, Istanbul, and Tokyo. Tehran, a city of roughly 13 million inhabitants, last had a major earthquake in 1840. Three active geological fault structures run through the urban area, and all of them are deemed capable of generating shallow-focus earthquakes of magnitude around 7. As about half of the city’s building stock is not antiseismic, the death toll in a future disaster has been estimated at between 400,000 and 3.4 million.
Gridding the globe with abstract lines has been put forward since the 2nd century BC, when Hipparchus suggested parallels of latitude and meridians of longitude constructed over a spherical model of Earth. Three hundred years later, Claudius Ptolemy refined these ideas and produced map projections with a graticule of longitude and latitude lines.

After the French-sponsored geodetic expeditions in Peru and Lapland in the 1730s, the ellipsoidal shape of the Earth became the basis for more accurate positioning of lines of latitude. After the International Meridian Conference of 1884 recognized the center of the transit instrument of the Royal Observatory at Greenwich, England, as the zero origin for longitude, the prime meridian, most nautical charts, and many maps agreed on the alignment of the longitude and latitude on small-scale charts and maps.

Spherical trigonometry is sufficient for distance and direction computations on the sphere, but the ellipsoidal shape of Earth makes these computations difficult. Additionally, the very meaning of distance is changed by different conceptions of appropriate paths from one point to another.

Gerardus Mercator designed a map projection in 1569 that allowed one to plot a single compass bearing as a straight line on a map. The course followed by steering this single azimuth is not the shortest path from one point to another, but often the efficiency of steering a single bearing and the ease of course design makes the longer path worthwhile. On a spherical Earth surface, the shortest path is a great circle. On the sphere, the meridians of longitude all describe great circles. The parallels of latitude describe small circles, not the shortest path between points of the same latitude, unless they are both on the equator. The shortest path over the surface of an ellipsoidal Earth is the ellipsoidal geodesic. Following any great circle or geodesic other than along a meridian or the equator requires continuously changing the heading with respect to true north. On the ellipsoidal Earth, spherical trigonometry is not sufficient for computations of distance, direction, area, path intersection, and other practical problems. Software to perform geodetic direction and distance algorithms can require several pages of source code. Software for projecting the ellipsoidal Earth onto a flat plane can be complex and slow to execute for large geodatabases.

There are other global coordinate systems based on longitude and latitude. In the plate carrée projection, the lines of longitude and latitude are mapped as though they were orthogonal and equally scaled everywhere. The World Geographic Reference System and Maidenhead Grid Squares divide the surface of Earth into nested rectangles in longitude, latitude space described by alphanumeric characters. These are useful for pointing to positions or small regions, but dividing the world up into equal areas using longitude and latitude is as impossible as flattening Earth with a single map projection without introducing distortions.

While longitude and latitude are useful, especially for storing of database vertices, for index maps, and for point descriptions, there are major problems for mapping, navigation, and spatial analysis. Hundreds of useful map projections have been devised, resulting in local and regional coordinate systems with which distances, directions, areas, and shapes can be portrayed and measured. These map projections, based on the notion of flattening all or a portion of Earth’s surface onto
a flat plane, all fail to portray the Earth without distortion. Cylindrical, conic, and flat planes are geometric surfaces on which features on Earth’s surface are projected. Cartographers can minimize one or more distortions of distance, directions, area, local shape, or global shape but never all of them. Small local regions can be portrayed on large-scale maps with reasonable fidelity, but tiling such maps to cover larger areas is not possible without overlaps or gaps. Global portrayals of Earth at smaller scales are often based on compromise projections that introduce minimal distortion of a few characteristics.

While Greenwich and the equator are meaningful concepts as the origin for longitude and latitude, they are not sufficient for mapping and positioning, especially at large scales. Geodetic datums are models of the size and shape of Earth and for the origin and orientation for longitude and latitude. The problem is that there have been many Earth shapes and many prime meridians, even those that pass through Greenwich. Distance and direction computations can differ significantly when different Earth radii are used. Maps and databases referenced to one geodetic datum can have point positions that differ by hundreds of meters when inappropriately tied to the wrong datum.

Geographic databases are often associated with a specific scale or resolution. They are appropriate for display and analysis within a limited range of scale. We live now in a time when large-scale mapping and spatial analysis is often done over long distances and wide areas. Research efforts are underway to develop multiscale data sets that can be used appropriately over a broad range of mapping and analysis scales. Scale-appropriate smoothing and thinning of large-scale databases could result in global databases based in longitude and latitude that could be projected on the fly for display and analysis. There are attempts to tile the globe and the ellipsoid with equal-area regions based on hierarchical global tessellations in triangular or hexagonal shapes. Global elevation data sets can be similarly processed for use at multiple scales.

Peter H. Dana

See also Coordinate Systems; Datums; Equator; Latitude; Longitude

Further Readings


projects in South America, Asia, and Africa. He is widely published in peer-reviewed professional journals and industry periodicals.

Eastman is one of the early pioneers of GIS. As an integrated GIS and image-processing system developed with the researcher in mind, Idrisi offers unique functionality not available in any other program.

Christopher D. Lippitt

See also Environmental Management; GIS in Land Use Management; GIS Software

Further Readings


ECOFEMINISM

Ecofeminism is the study of women’s connections to nature and how they inspire particular forms of environmental activism, stewardship, and spiritual attachments to Earth. This entry explores the evolution and key debates within ecofeminist thought, identifying some influential thinkers and suggesting avenues for further exploration. The three main themes within ecofeminism addressed here are (1) patriarchy and the domination of Earth, (2) Earth-centered spirituality, and (3) historical materialism and women’s knowledge.

The term ecofeminism often conjures up images of women chaining themselves to logging equipment, yet ecofeminism is far more than an activist and spiritual movement. Ecofeminism emerged as a field of study in the mid to late 1970s, heralded by the publication of a seminal paper by Sherry Ortner and a book by Susan Griffin, *Woman and Nature: The Roaring Inside Her*, both of which examined the associations of women with nature. These texts marked the beginning of more than three decades of feminist and other scholarship challenging the Enlightenment binaries through which modern Western culture has developed. While ecofeminism diversified and evolved, and alternative ecological feminisms and feminist political ecology emerged, the argument that the oppression of women and environmental degradation are connected remained an important contribution to knowledge. Ecofeminists assert that the association of women with nature and men with culture is not an innocent by-product of modernity but rather is integrally linked to a host of other binaries, including rational/emotional, nurturing/competitive, and nature/society.

Key to understanding the themes of ecofeminism is understanding what is meant by *essentialism* and the critiques of it. Essentialism refers to the idea that a particular group of people contain essential characteristics that give them privileged knowledge or physical attributes. Common examples of essentialist thinking are the assumptions that all women are inherently nurturing and that a particular race of people are inherently good athletes. Many ecofeminists embraced the idea that women are closer to nature and used it to insist that women are inherently better environmental protectors. Griffin’s work employed evocative and poetic literary devices to equate the violence done to the land with violence done to women. The phrase “roaring inside her” suggested that women, like nature, could fight back and were doing so, using their “nature”; here, it is assumed that patriarchy and women’s inherent nature lead to their understanding of and “rage” over environmental exploitation. But other ecofeminists have rejected this kind of essentialism and insist that it is women’s material existence or their experience that gives them an understanding of nature. They therefore postulate a *nonessential* association with nature. The tension between essentialist and nonessentialist renderings of the women-nature connection has in many respects been a defining feature of schisms within ecofeminist debates. This is elaborated as follows.

A core premise of ecofeminist thinking is that patriarchy and the domination of women
underlie other problematic social inequalities such as racism, heterosexism, and environmental exploitation. Ecofeminists have diverged over the implications of this association and the best way to overcome such binaries. Some have remained committed to the idea that women are closer to nature than are men and seek to celebrate that closeness and the privileged knowledge they believe it gives women. These so-called essentialist feminists have been widely criticized, although to suggest that their point of view has been rejected entirely is inaccurate. The second main focus of ecofeminist scholarship is that of spirituality. Spiritual ecofeminists seek to recover a metaphysical understanding of the world that worships Earth as a nurturing, living, female being to foster more sustainable environmental practices. Yet both an understanding of women as inherently closer to nature, and the notion of a female-centered spirituality, rely on essentialist notions of what it means to be a “woman.” Critics of essentialist thinking from within ecofeminism have emphasized that any privileged knowledge women have of nature arises from the material conditions of their lives. As women work in forests or fields or fulfill other (usually household-related) duties, they gain a deeper knowledge of environmental change and a better sense of environmental stewardship. It is their experience that makes them “closer to nature,” not their biology or inherent characteristics. Ecological feminisms and feminist political ecology have emerged from these critiques, although within ecofeminist thinking itself, a strong anti-essentialist focus continues to counterbalance essentialist thinking.

**Patriarchy and the Domination of Earth**

In the late 20th century, the feminist movement was deeply concerned with questions of patriarchy and its impact on a wide range of issues in society and scholarship. Especially important was Carolyn Merchant’s work, which explored the ways in which science and the Enlightenment transformed the ontology of “nature” from a living, maternal being into an inanimate entity governed by mechanistic laws based on Newtonian physics. Her overall argument was that our present overexploitation of nature is grounded in an ontology of the world that no longer values Earth as a living organism in its own right. Merchant asserted that the “living Earth” that was “killed” by science was female. The idea that Earth is female has not changed; the “death of nature” corresponded with the emergence of a patriarchal society that devalued and dominated women as well as nature.

Merchant’s work detailed examples of how society moved from worshipping and respecting “Mother Earth” in activities such as mining to an industrial, scientific understanding of geology that opened the door to far more damaging and exploitative forms of mining that are still seen today. She argued that the discourses and practices of domination were predicated on an equivalent devaluing and domination of women. Thus, the reversal of environmental destruction cannot occur without women’s emancipation.

Although Merchant’s work remains at its core anti-essentialist, it was an important inspiration for ecofeminists who failed to make distinctions among women. Merchant does not make fine distinctions between the kind of activities that might give women privileged knowledge of nature and those that might not. Rather, all “women” were assumed to have the same kind of sympathies and understandings of environmental change as a consequence of their close connection to nature, which has subjected Merchant to critiques of being essentialist. Ironically, some of the most important work in this vein has been by women from developing countries; yet it was women of color who later successfully challenged the dominant feminist thinking at the time, which failed to understand how race and class, along with gender, were equally important axes of oppression.

Vandana Shiva’s book *Staying Alive: Women, Ecology and Development* is perhaps the most famous exploration of the supposed innate connection women have with nature. Shiva drew from work on the grassroots Chipko environmental movement in the Indian Himalaya. Chipko activists fought commercial logging companies by chaining themselves to trees and sitting across roads to prevent logging from occurring. The movement gained global recognition in part because village women were on the front lines of such protests and in part because of their success at blocking logging companies on the ground. Shiva argued that the symbolic resources these
village women drew on to save their forests from commercial loggers were deeply rooted in their instinctive understanding of Earth as mothers and nurturers and spiritual ideas of women as protectors of Earth. From this position, she asserted that the inspiration for women risking their lives in front of logging machinery arose from these inherent connections to nature. In the context of modern Indian society, it is difficult to see women as powerful, but Shiva highlighted ancient stories and religious beliefs to show how women are in fact more powerful than men and certainly have a more profound understanding of environmental change. This work has been important in advocating for a global women’s movement, linked together by threats to women’s home environments. Shiva asserted that the inherent understanding women have of their environments would help bridge cultural gaps and provide a focus around which they could join together globally. Although Shiva’s work has been heavily criticized within India and by other feminists for inaccuracies and problematic assumptions about different women’s experiences, it has inspired women worldwide to defend environmental resources and was central in making “Chipko” a household name in many places. She also helped inspire others to explore situations where women were emerging as leaders in fighting environmental pollution or destruction.

Conceptualizing the relationship between women and nature as an essential one also challenged the hegemony of (male) scientific knowledge as the privileged source of information about nature. This helped open the door for alternative knowledges to inform debates about environmental change, and, indeed, traditional ecological knowledge (TEK) owes a large debt to ecofeminist work in this regard, even if it is rarely acknowledged. The TEK literature tends to suffer from its lack of attention to what feminist political ecologists call “gendered knowledges.”

**Spirituality and Restoring a Female-Centered Metaphysics**

One strand of ecofeminist thinking that has maintained an essentialist perspective on the relationship between women and nature is found in the writings on goddess worship and Earth-based spirituality. Vandana Shiva’s early work in many respects was part of this strand, but many ecofeminists in North America are also important contributors. Building from Merchant’s historical exploration, spiritual ecofeminists seek to recover the idea that Earth is alive and that human actions are threatening the health of this organism as much as they are threatening the individual species on Earth. James Lovelock’s ideas of Gaia have been taken up, and ecofeminists have a synergy with deep ecology in seeking to realign the human relationship with nature. This realignment can only take place when people embrace the importance of a spiritual connection with Earth and accept that every action affects the whole. It is therefore an ontology of holism, using ecological and organism metaphors to understand how humans affect the planet.

This kind of ecofeminist thinking has both been an inspiration and come under attack as it again reinforces essentialist ideas of women as well as nature and spirituality. It is unclear how this new Earth-centered spirituality sits in relation to the world’s dominant religions and, importantly, how it provides a better foundation for action than do the present ontologies.

**Historical Materialism and Women’s Knowledge**

Although celebrating the clear link between women and nature had some positive outcomes, essentialist conceptualizations of women rely on the notion of a fundamental female nature and ignore the diversity that exists between women. Such conceptualizations do violence to the very real differences between women, particularly the way race and class subject women, making the experience of being “a woman” far from universal. Women of color in the United States were some of the first feminists to challenge the white, middle-class bias in feminism. Latino American and African American women argued that in many contexts race was more salient in shaping their experiences of inequality, and thus the intersections of race, class, and gender need to be theorized.

Bina Agarwal used similar arguments to challenge Shiva’s essentialist rendering of the women and environment nexus. She argued that the relationship between women and their motivation to protect the environment was based on
material realities and not an inherent, close connection to nature. Drawing from her own work on women and fuelwood issues in the Himalayas, Agarwal demonstrated how many Indian women are responsible for the food and fuel needs of their families, requiring them to work closely with forests and agricultural fields every day. These activities give them an intimate knowledge of their ecosystems and a strong need to ensure that resources are used sustainably, as failure to do so results in increased work burdens for themselves. Agarwal named her feminist historical materialist account “ecological feminisms” to distance herself from the essentialist conceptions of women dominant in the literature of ecofeminism.

Agarwal’s appeal to historical materialism provided an empirical basis to the idea that women have unique environmental knowledge and, significantly, brought a political economic analysis into the debate around gender and environment. She argued that the material conditions of people’s lives are complicit in producing particular kinds of environmental problems and that these problems place extra burdens on women responsible for the subsistence needs of their families.

Although Agarwal sought to distance herself from ecofeminism, many ecofeminists have built on her ideas and rejected the early essentializing of Shiva and others. Rather than assuming that all women have the same experience, this work looked at how culturally specific ideas of gender define differences between men’s and women’s associations and knowledges in relation to the environment. The gender-environment nexus is therefore shaped by material practices, in particular men’s and women’s work practices, and culturally specific gender roles rather than innate understandings and spiritual connections. Many ecofeminists therefore argue that the emphasis needs to shift to how environmental issues affect men and women differently, without the assumption that women innately understand environmental issues better.

Andrea Joslyn Nightingale

See also Chipko Movement; Deep Ecology Movements; Feminist Environmentalism; Feminist Political Ecology; Gaia Theory; Gender and Nature; Indigenous Environmental Knowledge; Nature-Society Theory; Social Construction of Nature

Further Readings


ECOLOGICAL ECONOMICS

Goods and services from the natural world are vital components of the human economy. Ecosystem processes provide energy and regulate wastes, and natural resources are used for a variety of goods and services, including food, medicine, and recreation. Ecological economics is a transdisciplinary field that studies the allocation of natural resources, with emphasis on the view of the human economy as a subset of the ecological world. Drawing on expertise from the natural
and social sciences, physics, and other fields, ecological economists seek to include natural resources in the traditional economic view of capital and promote limitations of growth in favor of sustainable development. The field further redefines traditional ideas of sustainability to include responsible use of resources that does not preclude future generations from enjoying standards of living comparable with those of citizens living today.

The field’s conceptual basis was anticipated as early as the 1800s, when scholars began to express concern over limitless population growth and economic expansion. These ideas became joined with concepts from the physical sciences that highlighted energetic and material limits to growth to form discussions on resource use that include considerations from ecology, conservation biology, geography, and economics. To date, thousands of scientists and practitioners are organized in regional and international societies for ecological economics, and several degree programs have been established worldwide in ecological economics.

**A Break From Neoclassical Economics**

Although ecological economics shares many of the fundamental principles of neoclassical economics and environmental economics, its evolution represented a significant paradigm shift that drew natural resources into the concept of capital from which they had traditionally been excluded.

**Economy as a Subset of the Environment**

Ecological economics differentiates itself from neoclassical economics in its philosophical stance on the relationship between the environment and the economy. In the neoclassical purview, the economy exists as a system separate from its environment and with the potential for infinite expansion. It concentrates on the monetary exchanges among producers and consumers, whereby the optimal allocation of goods among alternative uses is determined primarily by market forces. Value can be generated irrespective of the natural resource base as long as beneficial technological change occurs rapidly enough.

In contrast, ecological economics applies laws from the physical sciences to present the human economy as a part contained within a larger ecosystem filled with finite resources. Growth of the economy necessarily depletes the base of natural resources, many of which are nonrenewable, and technology itself—however rapidly changing—cannot function without a minimum of materials and energy flow. In this view, economic growth is inherently limited. Thus, growth over time results in a shift from an “empty world” (i.e., many natural resources available) to a “full world” (i.e., most natural resources have been depleted or nearly depleted). During this transition, there is a point of economic growth beyond which human welfare is in fact reduced rather than increased, becoming uneconomic growth.

**Consideration of Natural Resources as Capital**

Ecological economics emerged in part from criticisms of the view of natural resources in traditional economics as fully substitutable with other standard input factors to production, such as capital and labor. In contrast, ecological economics stresses the relationship between resource flows and human-made production funds as complementary and not substitutable.

Ecological economists in the field have criticized the neoclassical economic production function

$$Q = F(K, L),$$

where the quantity ($Q$) of a good that firms choose to manufacture is a function ($F$) of capital ($K$) and labor ($L$) inputs to production.

Ecological economics modifies the production function to include explicitly nonrenewable and renewable resources ($N$ and $R$, respectively):

$$Q = F(K, L; N, R).$$

This modified equation accounts for natural resources as a limiting factor of production. Thermodynamic limits on material and energy use for specific technologies and for technical change
play a central role when specifying production functions and the underlying substitutability of factors of production for each other.

**Growth Versus Development: Themes and Policies in Ecological Economics**

Ecological economics views increasing welfare as dependent on the ability to make qualitative improvements versus the neoclassical emphasis on growth, which implies a quantitative increase in the flow of materials and energy from the ecosystem through the economy and back into the environment in the form of wastes. Development requires little or no additional throughput but emphasizes designs of existing commodities and institutions that improve the status of scale, allocation, and distribution of resources.

**Optimal Scale of the Economy**

In economic terms, *scale* refers to the volume of matter and energy used to provide goods and services in the economy. Many ecological economists posit that our current scale of natural resource use is unsustainable and that the human economy is approaching the full-world scenario, in which growth is uneconomic. Empirical assessments of many of the resources most critical to the economy suggest that continued expansion will come at very high costs. For example, 69% of the world’s primary commercial fish species are in decline, mineral resources are becoming scarce and declining in quality, and future projections of freshwater availability indicate that nearly 50% of the world’s population will experience water shortages in the next 50 years. This is in part due to negative externalities (e.g., pollution) resulting from economic growth and from failure of the market to capture scarcity of resources without monetary value.

Some ecological economists see the optimal scale of natural resource use as one that does not cause the resource in question to depreciate or erode over time but will provide a relatively constant stock for future generations. The challenge for natural and social scientists, as well as for policymakers, is to identify appropriate scales of consumption for various natural resources and to design incentives for firms not to extract these resources at higher rates.

Command-and-control regulations have historically been the primary mechanisms for setting limits on resource consumption. These policies set strict limits on pollution, extraction, or harvest levels and fine firms for violation. However, a recent shift in the discourse about the effectiveness of such policies favors more flexible solutions that provide incentives to reduce scale beyond one set cap in a more cost-effective manner. For example, “Pigouvian” taxes, named after the British welfare economist Arthur Pigou, involve charging firms per unit of pollution to achieve a socially optimal outcome. Similarly, Pigouvian subsidies may be levied to encourage the generation of certain benefits from production, for which there may otherwise be no monetary incentive. Tradable permits and quotas are also thought to provide monetary incentives for firms to reduce their scale of production and negative impacts on the environment to perhaps a greater extent than can be achieved through direct regulation.

**Efficient Resource Allocation**

Like neoclassical microeconomics, ecological economics is concerned with the allocation of goods and services (i.e., resource flow) among alternative uses, placing particular emphasis on scarce resources. While accepting the basic laws of supply and demand as useful tools of market analysis, ecological economics draws attention to the inefficiency of the market to allocate many types of natural resources due to various types of market failures. Open-access regimes such as fisheries, for example, are nonexcludable (i.e., traditionally lack property rights), and individuals may overexploit these resources because any costs incurred are shared among many other individuals also using the resource. Public goods (e.g., fresh water, clean air, scenic beauty) are subject to similar effects because they can be used for free and thus scarcity will not be adequately reflected in the market price. The production of goods can also produce positive or negative externalities (e.g., pollution)
that are not captured in the market price. Wherever these market failures occur, goods will not be produced or priced efficiently, contributing to the inefficient use of natural resources.

Environmental economics typically suggests that some of these problems can be solved by assigning prices to natural resources and taxing polluters. Ecological economics, while generally not averse to this solution, sees clear limitations and emphasizes the intrinsic and moral value of ecosystems and their role in supporting the human endeavor that is fundamentally different from the contributions that come from ordinary goods and services. Ecological economists typically suggest implementation of policy mechanisms that reflect the fact that natural goods and services do have some value without always attempting to calculate a dollar value for each.

**Equitable Resource Distribution**

Ecological economics places importance on the fair distribution of resources, wealth, and income for sustainability and responsible use of resources, even the fair distribution of resource flows between humans and other species. Income distribution can undermine sustainability because poor communities may not be able to afford to address environmental impacts, while the very rich consume vast amounts of limited resources and may be better able to adapt to growing resource constraints or decreasing environmental quality. There is also an ethical consideration of temporal distribution; future generations are viewed as having ownership of natural capital as well, and thus, their needs should be included in the picture of what is considered sustainable today. Some proponents of an equitable distribution argue that reductions in inequality of wealth would not only improve the condition of the ecosystem and its valuable resources but would have other positive social benefits, including economic stability and reduced incidence of crime.

Policies designed to redistribute wealth include progressive income and wealth taxes, mandated minimum wages, unemployment insurance, and welfare programs. Ecological economists also typically advocate expanded ownership opportunities as mechanisms to redistribute wealth. Employee Stock Ownership Programs and Consumer Stock Ownership Programs are designed to give stock and managerial influence to employees in firms so that profit-making decisions are spread over a larger number of individuals than those in the top tiers of management. This way, the external costs of profiteering may be to some extent internalized when a number of the invested workers live in an area influenced by negative impacts of production, for example. Trust funds can serve an analogous function for redistribution of natural capital, facilitating community ownerships of natural resources and associated profits.

**Questions for the Future**

With increasing populations and changing environmental conditions come lively debates about natural resource use and equitable distribution. In addition to broad concerns over environmental quality and demographic and temporal equity, specific recent concerns in ecological economics that will pose challenges to researchers in the future include the development of aquaculture versus open-sea fisheries; the ability of agricultural systems to meet food demand; the implications of genetically engineered seeds for food production; the availability and development of sinks for the vast amounts of waste generated by the human economy; the development of tools and institutions that help economies transition away from their current growth mentality; and the development of sustainable consumption habits.

_Rebecca R. Gasper and Matthias Ruth_

See also Common Pool Resources; Distribution of Resource Access; Ecological Footprint; Ecological Justice; Ecological Modernization; Energy and Human Ecology; Energy Resources; Environmental Ethics; Externalities; Neoliberal Environmental Policy; Nonrenewable Resources; Political Economy of Resources; Population and Land Degradation; Population, Environment, and Development; Renewable Resources; Resource Economics; Sustainability Science; Sustainable Agriculture; Sustainable Cities; Sustainable Development; Sustainable Development Alternatives; Sustainable Fisheries; Sustainable Forestry; Sustainable Production
The ecological fallacy arises when inferences are made about the characteristics of individuals in a population based on analyses using aggregate statistics of those same individuals. The general usage of the term applies to inferences made about individuals that are based solely on the group characteristics of that population. Such would be the case with stereotyping or geographical profiling, where individuals are characterized by the general traits of the area in which they live. In a more specific manner, the ecological fallacy occurs when an ecological correlation between variables based on statistical constants such as group rates or means is assumed to characterize individual members of that population. William Robinson first demonstrated that statistical correlations of variables aggregated by group or geographic area can differ markedly from correlations based on individual-level variables of that same population. The possibility of false or misleading associations has significant implications for geographical studies since relational inferences about members of a population are often based on data aggregated by spatial units.

In addition to the ecological fallacy, Hayward Alker identified other forms of fallacy that prevent valid generalizations from being made across different levels of analysis. The individualistic fallacy occurs when coarse-level aggregate relationships are inferred from individual or fine-level relationships. Cross-level fallacy exists when inferences are drawn from one subpopulation to another subpopulation at the same level of analysis. The universal fallacy applies when inferring generalized relationships based on nonrandom subsamples of a population. The selective fallacy typically occurs when a few cases are chosen to prove a universal point. Cross-sectional and longitudinal fallacies infer generalized relationships across many time frames based only on one point in time. Applying to all the above are contextual fallacies, where context or social structure may alter the strength or form of the causal or statistical relationship.

At the root of these fallacies is the concern that inferring generalized causal relationships based on spatially aggregated data can generate erroneous conclusions. Part of the problem is that errors arising from ecological inference are unquantifiable because of the lack of access to individual-level data. The spatial aggregation of data into defined areal units is common in geography and especially so with the increasing availability of digital data sets, digital boundary geographies, and geographic information systems. The Modifiable Areal Unit Problem is a derivative of ecological fallacy. There are many possible combinations of modifiable spatial units, such as voting districts, that can be used to spatially aggregate nonmodifiable units, such as the individual voter. As Openshaw demonstrated, aggregating data into arbitrary spatial units, which often have no meaningful geography, can greatly influence the outcomes of a study. Solving the ecological problem as to how to make inferences about individual behavior based on aggregate data continues to be an active focus of research in the social sciences.

Trevor M. Harris

See also Modifiable Areal Unit Problem
The ecological footprint is a measure of the biologically productive area that a human population requires to produce the resources it consumes, and absorb its wastes, using prevailing technology. A fundamental requirement for sustainability is that renewable resources must be used at a rate slower than the rate at which nature can replenish them and wastes be emitted at a rate slower than that at which the biosphere can absorb them. Societies that do not meet this minimum condition run ecological deficits.

To know whether humanity meets this requirement, and to properly manage ecological assets, measurement of the use of nature is required. Resource accounts are needed to track how much nature exists and how much is used by people. Ecological accounting operates like financial accounting: It tracks income (the ecological services that nature provides) and expenditures (human use of these ecological services). As with financial assets, it is possible to spend more than is being regenerated. But this is possible only for a short period of time. Such overspending depletes natural capital and cannot be sustained in the long term. Continued ecological deficit spending leads to environmental bankruptcy, eroded economies, decreased quality of life, and societal instability. In short, just as a successful business needs to keep track of revenues and expenditures, human society must keep robust accounts of demands on, and the renewal rates of, ecological assets. This is what the ecological footprint accounts offer.

The Ecological Footprint Analysis Method

The ecological footprint accounts measure people’s demand on nature. This demand includes both the resources that are consumed and the wastes produced. These resources are obtained from forests, cropland, fisheries, grazing land, and other ecosystems. The built environment compromises the land’s ability to provide biological resources. Additionally, ecosystems absorb and assimilate the wastes produced as a result of human resource consumption. The ecological footprint adds up these ecosystem services in terms of the biologically productive areas needed to provide the services. In other words, ecological footprint analysis builds on “mass flow balance,” and each flow is translated into the ecologically productive areas necessary to support these flows.

The ability of ecosystems to supply humans with natural resources is limited by climate, technology, environmental management practices, and the availability of water and solar energy. This supply of ecosystem services and natural resources is called biocapacity. When a population’s ecological footprint exceeds its biocapacity, ecological overshoot occurs.

The ecological footprints of nations can be calculated on an annual basis. Yearly figures are published by the nongovernmental organization (NGO) Global Footprint Network, along with methodological improvements that are used as reference points for applications by its partner organizations. These national accounts are presently the most scientifically scrutinized and well-documented national footprint assessments available. Adding up all nations’ footprints and biocapacities yields the global assessment. Overshoot measured on a global scale is an indicator of global unsustainability.

Global Footprint Network data show that humanity’s resource demands and waste production began to exceed Planet Earth’s ability to meet them around 1986 (Figure 1). By 2005, humanity exceeded the planet’s ability to provide biological resources by 31%—thereby dipping into the
natural capital stock. While the world average per capita biocapacity was 2.1 ha (hectares) per person, the world average per capita footprint was 2.7 ha per person. Ecological footprints vary significantly among nations. For example, the average footprint in Italy was 4.8 ha per person, against a domestic biocapacity of 1.2 ha per person, whereas the average footprint in Peru was 1.6 ha per person, against a domestic biocapacity of 4.0 ha per person (Table 1).

National Ecological Footprint accounts can also show local or regional ecological performance. For instance, ecological creditors (countries with a footprint smaller than their biocapacity) and ecological debtors (countries with a footprint exceeding their biocapacity) can be distinguished. In 2005, only 20% of the human population lived in ecological creditor countries. With increased resource scarcity, it may well be that geopolitical lines will shift from the division between “developing” and “developed” countries to that dividing ecological debtors and ecological creditors.

The ecological footprint measures one aspect of sustainability: the availability of, and human demand on, Earth’s regenerative capacity. A comprehensive assessment of sustainability requires other complementary measures as well: measures of social well-being, the depletion of nonrenewable resources, inherently unsustainable activities such as the release of persistent pollutants, and the degradation of ecosystems.

**History of the Concept**

The original ecological footprint methodology resulted from collaboration between Mathis Wackernagel and William Rees at the University of British Columbia in Vancouver. Their 1996 book made the concept more widely accessible.
<table>
<thead>
<tr>
<th></th>
<th>Population (millions)</th>
<th>Total Ecological Footprint (gha/pers)(^a)</th>
<th>Total Biocapacity (gha/pers)</th>
<th>Ecological Deficit or Reserve (gha/pers)</th>
</tr>
</thead>
<tbody>
<tr>
<td>World</td>
<td>6,476</td>
<td>2.7</td>
<td>2.1</td>
<td>(0.6)</td>
</tr>
<tr>
<td>High-income</td>
<td>972</td>
<td>6.4</td>
<td>3.7</td>
<td>(2.7)</td>
</tr>
<tr>
<td>Middle-income</td>
<td>3,098</td>
<td>2.2</td>
<td>2.2</td>
<td>(0.0)</td>
</tr>
<tr>
<td>Low-income</td>
<td>2,371</td>
<td>1.0</td>
<td>0.9</td>
<td>(0.1)</td>
</tr>
<tr>
<td>Africa</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Egypt</td>
<td>74</td>
<td>1.7</td>
<td>0.4</td>
<td>(1.3)</td>
</tr>
<tr>
<td>Ethiopia</td>
<td>77</td>
<td>1.4</td>
<td>1.0</td>
<td>(0.3)</td>
</tr>
<tr>
<td>Kenya</td>
<td>34</td>
<td>1.1</td>
<td>1.2</td>
<td>0.1</td>
</tr>
<tr>
<td>Madagascar</td>
<td>19</td>
<td>1.1</td>
<td>3.7</td>
<td>2.7</td>
</tr>
<tr>
<td>South Africa</td>
<td>47</td>
<td>2.1</td>
<td>2.2</td>
<td>0.1</td>
</tr>
<tr>
<td>Middle East</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Afghanistan</td>
<td>30</td>
<td>0.5</td>
<td>0.7</td>
<td>0.3</td>
</tr>
<tr>
<td>Iraq</td>
<td>29</td>
<td>1.3</td>
<td>0.3</td>
<td>(1.1)</td>
</tr>
<tr>
<td>Israel</td>
<td>7</td>
<td>4.8</td>
<td>0.4</td>
<td>(4.4)</td>
</tr>
<tr>
<td>Jordan</td>
<td>6</td>
<td>1.7</td>
<td>0.3</td>
<td>(1.4)</td>
</tr>
<tr>
<td>Asia-Pacific</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Australia</td>
<td>20</td>
<td>7.8</td>
<td>15.4</td>
<td>7.6</td>
</tr>
<tr>
<td>Bangladesh</td>
<td>142</td>
<td>0.6</td>
<td>0.3</td>
<td>(0.3)</td>
</tr>
<tr>
<td>China</td>
<td>1,323</td>
<td>2.1</td>
<td>0.9</td>
<td>(1.2)</td>
</tr>
<tr>
<td>India</td>
<td>1,103</td>
<td>0.9</td>
<td>0.4</td>
<td>(0.5)</td>
</tr>
<tr>
<td>Japan</td>
<td>128</td>
<td>4.9</td>
<td>0.6</td>
<td>(4.3)</td>
</tr>
<tr>
<td>Americas</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Argentina</td>
<td>39</td>
<td>2.5</td>
<td>8.1</td>
<td>5.7</td>
</tr>
<tr>
<td>Brazil</td>
<td>186</td>
<td>2.4</td>
<td>7.3</td>
<td>4.9</td>
</tr>
<tr>
<td>Canada</td>
<td>32</td>
<td>7.1</td>
<td>20.0</td>
<td>13.0</td>
</tr>
<tr>
<td>Colombia</td>
<td>46</td>
<td>1.8</td>
<td>3.9</td>
<td>2.1</td>
</tr>
<tr>
<td>Costa Rica</td>
<td>4</td>
<td>2.3</td>
<td>1.8</td>
<td>(0.4)</td>
</tr>
<tr>
<td>Cuba</td>
<td>11</td>
<td>1.8</td>
<td>1.1</td>
<td>(0.7)</td>
</tr>
<tr>
<td>Mexico</td>
<td>107</td>
<td>3.4</td>
<td>1.7</td>
<td>(1.7)</td>
</tr>
<tr>
<td>Peru</td>
<td>28</td>
<td>1.6</td>
<td>4.0</td>
<td>2.5</td>
</tr>
<tr>
<td>United States</td>
<td>298</td>
<td>9.4</td>
<td>5.0</td>
<td>(4.4)</td>
</tr>
<tr>
<td>Europe (EU)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>France</td>
<td>60</td>
<td>4.9</td>
<td>3.0</td>
<td>(1.9)</td>
</tr>
<tr>
<td>Germany</td>
<td>83</td>
<td>4.2</td>
<td>1.9</td>
<td>(2.3)</td>
</tr>
<tr>
<td>Italy</td>
<td>58</td>
<td>4.8</td>
<td>1.2</td>
<td>(3.5)</td>
</tr>
<tr>
<td>Poland</td>
<td>39</td>
<td>4.0</td>
<td>2.1</td>
<td>(1.9)</td>
</tr>
<tr>
<td>Spain</td>
<td>43</td>
<td>5.7</td>
<td>1.3</td>
<td>(4.4)</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>60</td>
<td>5.3</td>
<td>1.6</td>
<td>(3.7)</td>
</tr>
<tr>
<td>Europe (Non-EU)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Russia</td>
<td>143</td>
<td>3.7</td>
<td>8.1</td>
<td>4.4</td>
</tr>
<tr>
<td>Ukraine</td>
<td>46</td>
<td>2.7</td>
<td>2.4</td>
<td>(0.3)</td>
</tr>
</tbody>
</table>

**Table 1**  Ecological footprint and biocapacity results for selected countries (2005 data)

Since then, many others, including academics, leaders of NGOs, and business analysts have employed the concept and contributed to its methodological development.

To bring this community together, the Global Footprint Network was founded in 2003, with the goal of advancing the scientific rigor and practical application of the ecological footprint, and making the ecological footprint as prominent a metric as gross domestic product. The Global Footprint Network comprises a 23-member advisory board of leading scientists and politicians and has offices in Oakland, California (United States); Zurich; and Brussels. More than 100 organizations, spanning six continents, have become formal Global Footprint Network partners. The ecological footprint is now widely used by governments, communities, and businesses to set targets, monitor ecological performance, and evaluate risks and opportunities in an increasingly resource-constrained world.

The adoption of the ecological footprint as a trusted sustainability metric depends on the scientific integrity of the methodology, consistent and rigorous application of the methodology across analyses, and the clear and objective presentation of results. To meet these goals, Global Footprint Network and its partners have created a consensus-based committee process for improving the method and for developing international ecological footprint standards. In addition, a partner-based science committee oversees the development of the national footprint accounts.

Mathis Wackernagel and Kristin Kane

See also Carrying Capacity; Ecological Economics; Environmental Services; Sustainability Science; Sustainable Agriculture; Sustainable Cities; Sustainable Development; Sustainable Development Alternatives; Sustainable Fisheries; Sustainable Forestry; Sustainable Production

Ecological imaginaries are ways of conceiving nature with attendant but often unacknowledged social and cultural implications. These visions typically foreground purportedly neutral ideologies of nature while displacing often contentious political, social, and economic relations. Geographers seek to identify how seemingly durable conceptions of nature such as "rain forests" or relationships to nature such as "sustainability" attain their taken-for-granted status, that is, how they come to be accepted uncritically among large sections of the public. Articulated through material, social, and discursive practices, dominant ecological imaginaries reinforce ideologies of nature and society and naturalize particular conceptions of space, place, and landscape so that they appear inevitable, impartial, or cyclical. In short, they often function to make human relations with ecosystems appear to lie outside the capacity of people to change them. The term emerged from the common academic focus on meaning making and the consideration of nature in social theory and has been applied to purportedly apolitical concepts currently in vogue within urban planning, governance, redevelopment, and
environmental politics, as well as attempts to “green” capitalism in response to ecological crises. These linked ecological narratives frame policy discussions that are about more than just nature, throwing into question the neutrality of claims for ecological preservation and conservation. Instead of a singular nature, ecology, or public, investigators of environmental imaginaries argue for the recognition of multiple natures, ecologies, and publics but have been criticized for relying on explanations located primarily within discourse and not in material processes.

**Origins**

Firmly established within the “cultural turn” that has reshaped geography and the “green turn” in the social sciences, the term *environmental imaginaries* incorporates a recognition of the significance of ecological thinking to the constitution of place, space, and landscape as part of meaning making and identity. The attention paid to these ways of thinking or discourses was seen as a necessary corrective to the practice of geography as an abstract spatial science in which people and the way they make their worlds was of little importance to the discipline. Recognized as a constitutive element of social life, views of nature have become of increasing importance across the social sciences.

**Uses**

The consideration of ecological imaginaries has involved illuminating those representational strategies that naturalize particular conceptions of the landscape with social, economic, and political identities. It has also involved investigating environmental politics or the ecological imaginary of the “good city” in balance with nature and looking into taken-for-granted conceptions of environment, nature, conservation, and sustainability to see how social and economic visions are bound up with wish images expressed through representation of nature.

Ecological or organic metaphors underpin urban theories that explain political, economic, and social processes. For example, equating the city to the human body led to a concept of the role of transportation networks as circulatory systems, to arguments that parks need to act as the “lungs” of the city, and to a view of blight as cancerous decay. These narratives naturalized the processes that constituted urban landscapes so that they appear inevitable as opposed to the result of historical relations.

In other contexts, unpacking ecological imaginaries involves uncovering a “false” ecological view asserted for political or economic ends. For example, unpacking the cultural narratives around the American lawn requires unraveling the complex social, economic, and political relationships that have made this highly manicured, exotic, monoculture crop a ubiquitous part of the landscape and a central component of suburban identity.

Investigating environmental imaginaries illuminates representational and material spatial strategies that intertwine seemingly divergent concepts of nature, community, identity, and citizenship. Visions of improving on nature by controlling and harnessing rivers through waterworks and dams have reflected a desire to better society by supplying clean water to create a hygienic city and serve as a nationalistic demonstration of having attained the trappings of a modern city.

Recently, ecological imaginaries have sought to confront ecological and social dislocations that have resulted from modern urbanization and capitalist industrialization by mythologizing the control of nature and practice of urban design to suggest that environmental damage and a decline in sociability can be reversed. Livability and sustainability, considered as ecological imaginaries, mark conscious attempts to restore and reconnect ecological and social systems in compromised locations. The result is a politics of conservation, redevelopment, or aesthetic values that purports to be neutral but in doing so silences histories, voices, and opinions that fall outside the frame. This strategy has been pursued in the consideration of the epistemological politics of spatial thinking in relation to postcolonial critiques of contemporary environmental and resource politics as well as waterfront redevelopment projects.

The investigation of ecological imaginaries often highlights how utopian or dystopian themes become naturalized in ways that ignore the social relations that have constituted particular places, spaces, and landscapes. Instead of preserving or restoring “pristine” ecologies, this position argues
for more honest and transparent dialogues about the social and natural futures represented by ecological imaginaries. This shifts the focus of discussion toward considering what work the nature-based wish images do, what politics they support, which positions they erase, and ultimately what would constitute more just visions of social and natural landscapes.

Criticsms

By declaring all visions of ecological futures as “imaginaries” bound up with culture, society, and economy, it can become more difficult to make the sort of strong or objective claims about the exercising of power in the construction of landscape that make alternative spatial politics possible. Ecological imaginaries highlight complexity and context, which can obscure the attempt to criticize and oppose how such visions are mobilized in a political arena.

Chris Hagerman

See also Cultural Turn; Environmental Discourse; Environmental Imaginaries; Geographical Imagination; Nature-Society Theory; Political Ecology; Social Construction of Nature; Sustainability Science

Further Readings


ECOLOGICAL JUSTICE

Conceptions of justice typically address notions of what is right and what is fair. Ecological justice addresses how humans relate with nonhuman species and the natural world. Sometimes called *justice to nature*, it seeks to delineate human’s moral obligations to other species. This entry considers definitions, concepts, and issues central to ecological justice, which is different from *environmental justice*. The latter term refers to social justice environmentalism. Commentators attribute the term *ecological justice* to geographers Nicholas Low and Brendan Gleeson, but its scholarly antecedents date as far back as the mid 20th century, to Aldo Leopold’s “land ethic” and Rachel Carson’s *Silent Spring*, which sought to extend ethical behavior to biotic communities. As a research perspective, ecological justice recognizes that nature has intrinsic value and acknowledges the interconnections and mutual interdependence of all species. Proponents seek to expand the domain of “moral considerability” beyond humans to encompass animals, plants, and even inanimate objects such as rocks, rivers, and oceans.

There are several bases of ecological justice. Religious grounds often posit humans as custodians of the natural world and are founded on humans’ moral responsibility to other species, stemming from supernatural entities (God, Buddha, Allah, Australian Aboriginal Dreamtime beings, etc.). Instrumental grounds see current and future generations of humans as reliant on the natural world for their needs (e.g., food, medicine, and clothing). Without these species, humans may suffer extinction. More recently, some commentators have argued that humans have a kinship with nonhuman species, as fellow animals and “ecological citizens,” and are thus morally obliged to care for other species. In short, ecological justice seeks to reposition humans’ relationship with nature and to establish moral obligations to nonhuman entities.

Key Thinkers

There are too many key thinkers in this field to discuss in detail here. Some key contributions are outlined below, but others also merit attention.
(including those of J. Baird Callicott, Tom Regan, Ted Benton, and James Lovelock).

**Peter Singer: Animal Rights**

Perhaps best known for his book *Animal Liberation*, Singer argues for extending the notion of rights to animals, with some limits. He argues that many animals can feel pain, manipulate the world around them, have cultural expression, and develop deep familial bonds, thus making them worthy of ethical consideration. But Singer dismisses the notion of sentience (intelligence) as a test for moral considerability, arguing that very young humans or mentally disabled individuals may have lower intelligence levels than do many mammals.

**Christopher Stone: Do Trees Have Standing?**

Stone sought to challenge long-standing notions of legal inclusivity, expanding the community of legal right to include trees and inanimate objects. He proposed a system of legal rights for nature. Tried in U.S. courts, his propositions failed the test of jurisprudence but paved the way for legal recognitions of nonhuman species.

**Roderick Nash and Luc Ferry: Rights of Nature**

Roderick Nash traced the foundations of “natural rights,” exploring the philosophical and religious foundations of Western humanism. Together with philosophers such as Luc Ferry, Nash has researched how the Cartesian divide between humans and animals developed (i.e., a worldview that separates nature from culture) and how various religious and philosophical traditions have allowed for the development of ecological ethics (e.g., the animal trails of the Middle Ages). Ferry and Nash have challenged some ideas within the deep-ecology movement as being potentially fascist (e.g., what is natural, and place-based ecological connections that determine who and what belong where).

**Arne Næss: Deep Ecology**

Moving beyond Singer, Arne Næss has proposed a doctrine of equal rights of nature, espousing that all living entities are inherently valuable. Deep-ecology proponents hold nature to be sacrosanct and seek to protect “pristine nature” from human-induced harm by cultivating ethical, spiritual, and emotional identification with nature. But the ecocentric ethics of deep ecologists have been criticized for rendering the idea of rights meaningless, for not paying enough attention to social inequalities, for essentializing nature (ignoring the inherent differences between humans and nonhuman life forms), and for risking social Darwinism (i.e., naturalizing famine or legitimizing paternalism).

**Val Plumwood: Ecofeminism**

The late Val Plumwood strove to develop a feminist basis for ecological justice. She and other ecofeminists have sought to understand and undo structures of domination and oppression (e.g., class, patriarchy, racism, and speciesism) within the broader goal of theorizing how humans relate to nature and how in turn these relationships shape various power relations. They particularly challenge the association of women and nature—and women’s supposedly closer affinity to nature, defined by their “natural” reproductive capacities. This thinking they term *mutual inferiorization*. The ecofeminist project is framed around the liberation of the other (including both women and nature) from masculine, rationalist oppression.

**Geographers’ Contributions**

Within geography, ecological justice proponents include ecosocialists, ecofeminists, human ecologists, and animal geographers. Two areas merit closer attention—animal geography and political ecology.

**Animal Geographies**

Geographers such as Jennifer Wolch have over the past decade radically redefined the field of human geography to encompass the study of animals and nonhuman nature. Sarah Whatmore, Jennifer Wolch, and Bruce Braun—among others—have sought to create a “more than
human” geography, studying for instance how humans are already “hybridized” with nature (e.g., cities are socionatural entities), how human-animal relations have shaped spaces and places (e.g., how ideologically charged animal representations are used to construct ethnoracial identities), and how linkages between ideas about animals and racism, sexism, colonialism, and nationalism have configured cultural practices and power relations. Animal geography has often entailed reconceptualizing animals as what Sarah Whatmore has called *strange people*, showing how animals possess agency and subjectivity and are worthy of moral consideration and ethical treatment.

**Political Ecology**

Recognizing the interconnections between environmental degradation and socioeconomic exploitation, political ecologists have shown how the workings of capitalism (re)produce large-scale human and animal suffering and environmental harm. The political ecology perspective has cast light on how nature is distributed within cities; how cities function as metabolic entities and as habitats for myriad species; how capitalism “produces” nature—differentially shaping access to nature and determining who and what lives and dies; and how habitat destruction is linked with consumptive lifestyles (e.g., lawn chemicals and groundwater pollution). Understanding political and economic processes and institutions reveals impediments to the flourishing of humans and nonhuman nature.

Animal geographers and political ecologists have not only broadened our understanding of nature-society relations but also sketched out research agendas that place ecological justice as a subject of inquiry and goal of action-oriented research.

*Jason Byrne*

---

**Further Readings**


---

**ECOLOGICAL MAPPING**

Ecological maps are playing an increasing role in land management and planning. Their purpose is to display units of land of various sizes that reflect differences in response to management and resource production capability. An ecological map shows an area divided into ecosystems—that is, areas within which there are associations of interacting biotic and abiotic features. How these features are associated or integrated can be shown at two general levels. One level shows the integration within the local level, and another shows how the local area is integrated and linked with other areas across the landscape to form larger systems. All these areas are ecosystems, albeit at different scales or relative sizes.
Subdivision of land into systems of different sizes is needed for several reasons. Because of the linkages between systems, a modification of one system may affect the operation of surrounding systems. Furthermore, how a system will respond to management is partially determined by relationships with the surrounding systems linked in terms of runoff, groundwater movement, and microclimate influences. Because ecosystems are spatially nested in each other, each level subsumes the environment of the system at the level below it. Therefore, it conditions or controls the behavior of the system at the level below it. Understanding these relationships is important in analyzing cumulative effects of action at one scale and its effects at another.

Given this need, landscape ecologists and ecosystem geographers have been interested in hierarchical schemes of ecosystem units. The nomenclature and number of levels in these schemes vary. One scheme that has been adopted by the U.S. Forest Service and others for use in ecosystem management recognizes ecosystems at three scales of perception. The smallest, or microscale, ecosystems are the homogeneous sites commonly recognized by foresters and range scientists. At the mesoscale, linked sites create a landscape mosaic that looks like a patchwork. At the macroscale, mosaics are connected to form larger systems. These units of connected mosaics are called ecoregions.

The fundamental question facing all ecological land mappers is “How are the boundaries of different-sized systems to be determined?” Different methods have been used to identify units where ecosystem components are integrated in a similar way, thereby classifying land as ecosystems. One method is to overlay maps of ecosystem components at a similar resolution to look for associations of these components to identify ecological units with similar patterns. Some ecosystem mappers have employed a gestalt approach in which boundaries are drawn intuitively around areas that appear homogeneous. Others used multivariate clustering to classify grid cells. A map is produced by drawing lines around cells of similar class. These methods are termed empirical because they are descriptive and do not explain why units should be distinguished.

In an alternative to the empirical approach, a major emphasis is on mechanisms that produce the pattern of ecosystem distribution. Establishing a hierarchy of ecosystem boundaries is based on an understanding of the formative processes that operate to differentiate the landscape into ecosystems at various scales. The units derived from such an approach are termed genetic in that they are predicated on an understanding of the causal processes that control the pattern of ecosystems. Understanding spatial relationships between causal mechanisms and resultant patterns is key to understanding how ecosystems respond to management.

The notion that knowledge of processes should define landscapes has a long history in geography (cf. William Morris Davis). Building on this precedent, several workers have outlined principles to study and delineate ecosystems using the genetic approach. Delineating units involves analyzing factors controlling the geography of ecosystems at various scales and then using significant changes in controls as boundaries. A spatial hierarchy is constructed by successive subdivision of large ecosystems on the basis of controlling factors operating at different scales. Proponents of this approach argue that because subsystems can be understood only within the context of the whole, a classification of ecosystems begins with the largest units and successively subdivides them. This approach “from the top down” is advantageous as it provides a meaningful global context for local concepts and problems.

A primary value of the genetic or controlling factor approach to ecosystem mapping is that it screens out the effects of disturbance or plant succession because boundaries are not based on existing conditions. By this approach, permanent boundaries can be identified that allow ecosystems to be recognized regardless of current land cover or condition.

This approach is based on understanding the role of climate in ecosystem differentiation. To a considerable extent, climatic factors determine the boundaries of ecosystems at all scales. The most important of these is climatic regime, defined as the diurnal and seasonal fluxes of energy and moisture. As these fluxes change, the kinds and patterns of dominant life forms of
plants and animals change, as do the kinds of soils. As a result, ecosystems of different climates differ significantly. Controls over the climatic effect change with scale. Understanding these controlling factors on a scale-related basis is key to setting ecosystem boundaries.

At the macroscale, the ecosystem patterns are controlled by the macroclimate, which is the climate that lies just beyond the local modifying irregularities of landform and vegetation. Several factors are involved:

**Latitude:** If Earth had a homogeneous surface, circumferential zones of climate would result from variations in solar radiation and resulting atmospheric circulation. These zones would be divided along lines of latitude.

**Continental position:** Because Earth’s surface is divided between land and sea, each with quite different thermal characteristics, the boundaries of climatic zones are “bent” as they cross the continents. These zones also fluctuate annually due to the inclination of Earth’s rotational axis.

**Elevation:** Earth’s landmasses are not, however, of a low, uniform elevation. The geology of the landmasses creates topographic variation. This topographic variation further affects the climate and therefore the regional distribution of ecosystems.

At the mesoscale, within the same macroclimate, broadscale landform patterns (topography and geology) break up the zonal pattern and provide a basis for further delineation of ecosystems. For example, geologic structure comes into play, creating mountains and valleys that create another climatic gradient: a combination of orographic precipitation and rain shadow.

At the microscale, the landforms created by crustal activity are shaped by processes and agents of denudation, creating a group of sequential landforms, such as dissected topography, where taller, denser vegetation demarcates the north-facing slopes. The local variations in topography will cause small-scale variations in the amount of solar radiation received, create topoclimates, and affect the amount and distribution of soil moisture. The variables will subsequently affect the biota, creating ecosystem sites as subdivisions of the larger zone.

To summarize, all natural ecosystems are recognized by differences in climatic regime. Regional ecosystems (ecoregions) are areas of a homogeneous macroclimate. Landform is an important criterion for recognizing smaller divisions within macroecosystems since it exerts the major control over the climate (through aspect and steepness of slope) at the meso- and microlevels. As modified by landform, climate offers the logical basis for delineating ecosystems, both large and small.

Boundaries are recognized by perceived changes in the ecological relationships of climate, vegetation, and landform. But when vegetation has been drastically disturbed or even destroyed, boundaries between potential ecosystems can still be mapped to coincide with changes in those landform characteristics known to regulate the reception and retention of energy and moisture.

Robert G. Bailey

See also Biogeography; Ecoregions; Ecosystems; Environmental Management

Further Readings


ECOLOGICAL MODERNIZATION

Ecological modernization is shorthand for two ideas: (1) that it is possible to maintain or increase the rate of economic growth and protect the environment and (2) that diseconomies and ecological harm may be diminished by policy correctives and technological fixes that design environmental criteria into economic systems. Its apparent appeal lies in its capacity to generate positive-sum solutions to problems conceived as zero sum, move beyond remedial and regulatory environmental strategies of the 1970s, avoid structural change seen as intractably difficult, and accommodate (however uncomfortably) both radical environmental critiques and neoliberal economic practices. In short, the term refers to the restructuring of the capitalist economy along environmentally sound lines. Nevertheless, it has been criticized for perpetuating social injustices, economic unfairness, and environmental harm because it remains inside the capitalist system, from which stem many of the problems of modernization.

Modernization and Globalization

Modernization is a term and idea describing various pathways for human and social development and various changes in social and spatial relations over time. These processes involve modifications to production and consumption, as well as adjustments to industrial practices, land use, migration, settlement, transportation and social, economic, and political organization. Modernization and globalization are interrelated, the latter enacted or operating at various spatial scales. It gives effect to increased and accelerated flows of financial and other transactions, capital, resources, goods and services, ideas, people, or communications. These flows are unevenly distributed, with varying consequences, both positive and negative.

Concerns about the harmful effects of modernization and globalization have given rise to diverse environmental values. Emphasis is sometimes placed on the intrinsic or essential worth of nature or the environment. Sometimes the instrumental or practical worth of such entities for human needs and desires is stressed. On balance, however, whether environmental values arise from self-interest or selflessness, they prompt calls for nature or the environment to be better conserved and managed, given the growth and globalization of modernization’s damaging effects. Among such effects are poverty, malnutrition and ill health, excessive consumption and the unfair distribution of goods and services, anthropogenic or human-induced climate change, and habitat and species loss. Such loss occurs across all habitat types at all latitudes and arises from inappropriate forms of urbanization, primary and secondary production, and other human activities and from the pollution of the environment by hazardous substances.

Sustainable development is one prominent response to such effects, gaining rapid authority in international governmental circles from the mid 1980s via work by the World Commission on Environment and Development (WCED). The WCED described sustainable development as development that meets the needs of the present without compromising the ability of future generations to meet their own needs, stressing the crucial need to ease poverty and work within social and technological limits. The idea began to displace the established, if controversial, agenda that questioned unfettered economic growth (free market capitalism), and sought to promote local economic self-sufficiency and a steady-state approach. Among the chief advocates of such economic reform, the ecological economist Herman Daly suggested that the modern “evils of growthmania”—a prevailing attitude that there is no such thing as enough—cause social injustice, economic malfunction, and environmental harm.

Conversely, members of the WCED proposed that human development needs could be met within existing economic systems. This idea was entrenched via major United Nations conferences on environment and development, which gave rise to Agenda 21 in 1992 and Action 21 in 2002. Arthur Mol and David Sonnenfeld have suggested that these international events served to codify how changes to economic, social, and
environmental relations could be undertaken inside modernization and globalization processes.

The simultaneous rise of neoliberal forms of governance during the same period goes some way to explain the success of this codification and its use in government policy. Neoliberalism, like sustainable development, is informed by challenging questions about how to live that are linked to social, economic, and ecological values and systems. Neoliberalism includes a range of philosophical and practical developments of the traditional classical liberal agenda. Classical liberals argue that no other liberties are possible without a guarantee that benefits will flow from property and freedom of contract—hence their backing of laissez-faire capitalism and the unregulated exercise of choice in the marketplace. For social liberals, social equality is more important than economic freedom, and although the latter remains important to them, they argue that markets must meet the basic needs of all, including their social, political, and economic goals.

Neoliberalism has several variants, but these hold in common the following: fiscal restraint to create surplus budgets; free trade; privatization of publicly owned goods, services, and infrastructure; legal security for property rights; minimal government intervention; and deregulation of labor and financial markets and (by extension) of laws and rules for environmental management. Like sustainable development, neoliberalism has profoundly affected general understandings of the conception of society, the place of the citizen, and the role of government. Among these new understandings is the idea of responsible autonomy, that citizens are accountable, dependable, and conscientious, act in their own (enlightened) self-interest, and do not depend on government for their welfare.

Given its apparent accord with neoliberalism and its capacity to shore up “business as usual,” the rapidity with which sustainable development was mobilized in national and international systems of governance is not surprising; it has served, intentionally or otherwise, to protect vested interests and the status quo. While criticism of the concept has been trenchant, it remains remarkably resilient and continues to have widespread influence. Under the influence of neoliberalism, for some analysts sustainable development represents the “greening” of modernization and has given rise to ecological modernization theory (EMT).

**Ecological Modernization Theory**

Much of the work on EMT originates from Germany, the Netherlands, and the United Kingdom, with empirical studies emanating from other parts of Europe and North America; Australians have also been instrumental in advancing significant work in the field. Chronological, national, and theoretical divergence is inherent to EMT, but there are several identifiable stages to it. Initially, EMT was championed by the two German sociologists, Joseph Huber and Martin Jänicke. Huber sought to justify a shift from sufficiency to efficiency and to explain how technological innovations could help environmental reform, particularly in industrial production, especially if the free market were enabled and bureaucratic strictures to its rationality removed. Jänicke’s early work, focused on environmental policy, also emphasized what Maarten Hajer later described as a *techno-corporatist* form of ecological modernization centered on industrial ecology and economic and technological efficiency.

Numerous critiques of such techno-corporatist works exist, many by those who had earlier promulgated ideas about sufficiency, by those otherwise deemed politically “red-green,” or by those seeking to refine EMT from the “inside.” Arthur Mol and Gert Spaargaren are among the last group, and they have argued that critiques of capitalism have been important in improving EMT, having generated at least three new insights. The first is that capitalism is responsive to environmental change. The second is that environmentally sound production and consumption are possible but require diverse and targeted environmental reforms. The last stems from an assertion that all chief alternatives to the existing economic order are infeasible and assumes the necessity of transforming free-market capitalism while protecting society’s sustenance base—sustainable development recast.

A second identifiable stage in EMT produced work focused more on the institutional and cultural dynamics of ecological modernization and on national and comparative studies of industrial production. Maarten Hajer has described this
stage as reflexive ecological modernization. His work, an early example of the empirical testing of ecological modernization through policy processes, maps a shift in EMT’s focus away from understandings of nature as a free good and a sink for wastes and toward anticipatory techniques of environmental policy making, new roles for science in policy, and new legislative and regulatory frameworks emphasizing risk and uncertainty.

The Risk Society and Reflexive Modernization

Work theorizing the inherent risk of life, and of the uncertainty that typifies it, emerged at the same time as that on ecological modernization. Ulrich Beck was among the first to posit that ours is a risk society, involving a shift from modernization as the quest for wealth using industrialization to master external threats (often sourced in the environment) to modernization as a way to manage the inherent dangers of those very processes of industrialization. A consequence of this shift is that modern environmental risks are understood to be globally democratic insofar as the uncertain effects of threats such as climate change, pesticide residues, or epidemic diseases do not (fully) recognize class differences, and therefore, such differences are not able to explain the distribution of risk among the population.

Ulrich Beck later developed these ideas as reflexive modernization, in which some commentators now position ecological modernization. He argued that modernization actually and paradoxically dissolves the foundations of modern industrial society, a crucial point that highlights the inadequacy of the central institutions of that society to solve the ecological crisis using existing means; these institutions are, then, both flawed and incomplete. For scholars such as Maarten Hajer, who posits a new reflexive ecological modernization, the challenge now lies in finding new institutional arrangements and in correcting the bias that favors particular forms of economic and scientific knowledge and practice that have proven destructive and incapable of addressing such tendencies. This challenge is pressing.

Elaine Stratford

See also Ecological Footprint; Ecological Justice; Environmental Ethics; Externalities; Globalization; Modernity; Modernization Theory; Nature-Society Theory; Neoliberal Environmental Policy; Neoliberalism; Nonrenewable Resources; Political Economy of Resources; Population and Land Degradation; Population, Environment, and Development; Renewable Resources; Resource Economics; Sustainability Science; Sustainable Agriculture; Sustainable Cities; Sustainable Development; Sustainable Development Alternatives; Sustainable Fisheries; Sustainable Forestry; Sustainable Production

Further Readings

ECOLOGICAL REGIMES

Ecological regimes can be defined as the way societies, institutions, or groups manage relations to their physical environment. This notion designates fields of activity, which are as such normatively regulated and which thereby determine relations between social and physical structures.

Regulating Society: Environment Relations

Since physical interventions in the world are often constitutive elements of actions, the coordination of actions or interactions incorporates physical aspects into social order at any scale. The regulation intends to establish or maintain specific environmental conditions. Ecological regimes thus constitute a realm of objects and related targets.

Typical examples of ecological regimes are the regulation of forestry, agriculture or other forms of land use; taking precautions against epidemics, floods, avalanches, and other natural hazards; energy policies; or the waste systems. Since literally everything may become involved in actions or interactions, there is no a priori limitation to the object realm of ecological regimes. However, the regulative targets determine what belongs to a regime or not. Insofar as ecological regimes are constituted by a normative coordination of actions, they resemble institutions. They may encompass institutions as well as individual actions, though.

Toward an Integration of Social and Physical Structures

That people do not make their history under self-chosen conditions is obvious. What is controversial, however, is what ways and to what extent material nature and artifacts contribute to the constitution of society. In general, the social sciences, including human geography, obey the rule of explaining social facts by referring to other social facts and processes exclusively. The purpose of this methodological imperative is to ensure that research avoids any kind of physical determinism, such as environmental determinism, normative naturalism, or racism. However, approaches such as ecological economics, science and technology studies, environmental philosophy, actor-network theory, ecofeminism, or political ecology strive to integrate material elements into their conceptions of the social. In this context, the concept of ecological regimes offers an approach to study the interrelations between social and physical structures that avoids physical determinisms and does not make the social an absolute.

By producing and referring to physical facts, actions and interactions may link social conditions to physical conditions. Consequently, the latter may even steer social processes to some degree. If, for instance, an annual budget for road repairs depends on weather conditions, then physical conditions are allowed to determine that money is spent for certain purposes instead of others.

The concept of ecological regimes might play a key role in developing perspectives in social theory that stress the coevolution of “nature” and “culture.” In contrast to notions of hybridity, which tend to sweep the specific differences between human actors, animals, plants, and physical conditions generously away, analyses setting out from ecological regimes trace the integration of physical and social facts down to individual actions or interactions and their regulation.

There are related notions in use. Like Christoph Görg, several authors address the regulation of society-nature relations, and when referring to
the shift from wood to coal or oil as the primary energy resource, Marina Fischer-Kowalski and Rolf Peter Sieferle speak of different socio-ecological regimes. For geography, this concept offers not only possibilities to relate physical and human geography in novel ways but also an alternative to the often problematic use of “space” or “landscape” as concepts, which somehow should integrate material and social facts.

**Empirical Analysis of Ecological Regimes**

The empirical analysis of ecological regimes preferably sets out from the regulation of certain types of actions. Figure 1 illustrates how certain actions establish relations to the social and the physical spheres and thereby integrate material circumstances into social structures. The scheme may be read as an abstraction of ecological regimes, and it may serve as a guideline for empirical analysis. In addition, Figure 2 illustrates how physical facts enter the coordination of actions. Unlike traditional social sciences, the concept of ecological regimes does not reduce the physical environment to a set of objects that play only a passive role within processes of social structuration. Rather, insofar as actions refer to and intervene into the physical sphere, physical facts become constitutive elements for actions and for the coordination of actions. Following Bruno Latour, one may conceive of interactions as an interplay of intersubjectivity and interobjectivity.

Like institutions, ecological regimes are reproduced or altered through discourses, which challenge or defend the legitimacy of the objectives and the means of regulation. As long as physical
conditions are not ignored, the empirical investigation of ecological regimes may also be carried out by discourse analyses.

Wolfgang Zierhofer

See also Actor-Network Theory; Coupled Human and Natural Systems; Ecological Footprint; Ecological Justice; Ecological Modernization; Environmental Ethics; Environmental Management; Human Ecology; Nature-Society Theory; Neoliberal Environmental Policy; Political Economy of Resources; Population and Land Degradation; Population, Environment, and Development; Sustainability Science; Sustainable Development; Sustainable Development Alternatives

Further Readings


Ecological risk analysis (ERA) is a method used to analyze and describe the environmental impacts of a project, program, or policy with the aim of parameterizing the cause-effects chain in a structured but overall simplistic way. In most applications, ERA is directly combined with environmental risk assessment (ERAS). In ERAS, the social dimension about the ecological value of the system parameters or indicators is added to the analysis in terms such as good, bad, action needed, low, high, and so on. The basis of the analysis is the best but incomplete and uncertain environmental information. The problem of missing knowledge and local information about the interactions in ecological systems and the different timescales of environmental systems must be solved by qualitative approximations and knowledge. ERA results only in assessments about the system behavior, and it is not suitable for the prognosis of landscape developments.

ERA is a method of impact analysis, for example, part of an ERAS procedure for usage in
Environmental Risk

In ERA, the term *environmental risk* is defined as the degree of interferences of the natural resources caused by significant changes in a planned project. This risk definition differs from the usage of the term *risk* in decision theory, which focuses on probabilities of risks. An ERA application in planning normally results in assessment values of a risk index by aggregation of different aspects, parameters, or indicators (e.g., soil, water, climate, fauna, flora, human, culture, heritage, social and economical conditions and their interferences) of the environment. The main aspect is the differentiation between analysis or modeling and the valuation step. The analysis and modeling normally result in nominal, ordinal, or cardinal outcomes (e.g., a soil erosion risk does [not] exist, the soil erosion risk is in the class between 1 and 2 tons ha\(^{-1}\) yr\(^{-1}\) [per hectare per year], or the soil erosion risk will be 11.5 tons ha\(^{-1}\) yr\(^{-1}\)). The assessment categorizes the results of the environmental analysis or modeling to a societal view of risk, for example, the result of an analysis of 11.5 tons soil erosion per hectare per year is compared with thresholds, norms, qualitative goals, or standards given by law or society by categorizing the cardinal measurements or modeling results into a class of high erosion risk. The categorization of the assessment into ordinal classes of risks is a standard application in environmental risk analysis.

ERA is used for different applications in practice. When focusing on sectoral approaches in agriculture, forestry, industry, tourism, or urban development, ERA applications use sectoral indicators sets (e.g., for the assessment of the sustainability of the land use of farms). When analyzing, for example, the impacts of a planned infrastructure project, a wide range of potential impacts on the environment is within the scope of the investigation.

The aim of ERA in planning is the evaluation or assessment of the suitability of a planned use change or use variants to locate a land use change in the best place. The operationalization of the technique can also be done if essential information is lacking by using proxies for the analysis of indicators or causal chains. The approach is based on three steps, by describing (1) the sensitivity of the landscape, (2) the intensity of the potential impacts of the planned land use, (3) and the impairment of landscape sensitivity as a result of the risks associated with the project.

ERA for a planned project to describe the sensitivity of the landscape is applied in several logical steps. The situation is analyzed first without the impacts of a project—that is, a baseline is estimated, which is a function of the sensitivity of the geographic area. The impacts on potential natural factors are then analyzed by using the knowledge of several disciplines about a single factor. The following descriptors are analyzed: the impacted natural factors; the services provided by the natural factors; the indication of their suitability for different land uses (ecological functions, goods, and services); and the ecological interdependencies in general, also by including the indication of their suitable usages and ecological interdependencies in a spatially explicit way for the area of the planned project. An additional step combines thresholds and the analysis with the resulting maps of sensitivity of the natural factors. The procedure is applied for each factor, and maps of classes of sensitivity are produced. A formal approach for the analysis might use a decision tree for each natural factor or a specialized and detailed modeling of the factors.
The intensity of the potential impact of the planned land use is analyzed as follows. The environmental impacts of the planned project are formalized by starting with a description of the planned land usages and an analysis of their impacts on the natural factors. Following this step, indicators are chosen to describe the impacts of a natural factor, which is generally followed by a spatially explicit description of the impacts by the indicators. Analysis of the intensity of the planned land use impairment results in a set of maps of the factors chosen in classes of impact intensity. Again, the formal approach for the analysis is the use of a decision tree for each natural factor or for specialized and detailed modeling of the factors.

The result of ERA is a determination of the impairment or sensitivity of the landscape based on impacts of the planned project in terms of risks of ecological impacts. For this purpose, steps of sensitivity analysis and impact analysis are merged with the help of a combination matrix when combining the maps that include the classes of sensitivity and the maps of impact intensity into classes of risk for each factor. In modern applications, ERA results in maps for each single factor. The resulting table of an ERA in EIA contains the project’s impacts on the different factors included; a further aggregation, for example, into a figure representing the total impact of the project will not be done because of the difficulty in aggregating different factors into a commensurable sum or index. Decision makers in planning will understand the ERA results better if impacts of a project in the form of each single factor are demonstrated and, furthermore, a decision on the basis of an ERA is not predetermined by the ERA study. The decision will be made by competent authorities or by politics. Owing to the reasons mentioned above, ERA is the standard method for EIA applications and is often compared with “utility value analysis” methods. The latter method has been used recently only for analyses of very large and complex projects when different time steps are analyzed (e.g., to assess the impacts of an open-cast lignite mine, after which a lake is planned, and the impacts on groundwater that will result from changes during and after the project period).

The mapping, analysis, and modeling in ERA are done today using many different approaches, including geographic information systems (GIS). Only the main aspects of these applications are briefly described here because of the rapid progress in their availability; the processing of new data, especially about climate change; and land use on the basis of remote sensing technologies. In early applications of ERA, GIS has been used only for the production of maps. Modern systems-oriented approaches of ERA use a multiplicity of digital data layers in the GIS, as an analysis tool, to model the sensitivities of the landscape and the project’s impacts. Because of the diversity of landscapes, projects, and the projects’ impacts, each ERA application will use a combination of suitable tools or models for each factor.

When digital data are publicly available, the scientific and pragmatic aspects of the methods should be applicable to analysis or modeling of the project’s risk, which is the focus of interest of ERA. Several frameworks for ERA are available in practice (e.g., at the U.S. Environmental Protection Agency or the European Commission). In geography, several assessment models have been developed for single factors or combinations of factors. Examples include SWAT (Soil and Water Assessment Tool) and MULBO (Multicriteria Landscape Assessment and Optimisation) and tools for multicriteria landscape assessment and optimization. SWAT is a river basin–scale model developed to quantify the impact of land management practices in large, complex watersheds. The model is physically based and computationally efficient, uses readily available inputs, and enables users to study long-term impacts. MULBO offers an open and structured framework to analyze and assess the risk surfaces of landscape functions by GIS. The optimal distribution of land uses are modeled in MULBO on the basis of linear programming to find ideal compromises between multiple goals on the basis of landscape functions and project goals.

An increasing number of modules for applications are available for more detailed analysis. A new focus should be to calculate the data
uncertainties when using GIS-based approaches and when combining data from different sources. The intersections and unions of the data layers can sometimes result in misleading outcomes.

**Further Developments**

New developments in policy analysis and stakeholder participation open new fields of application for ERA. When widening the focus of EIA from the analysis of natural factors to a general analysis and assessment of sustainability, the spatially explicit assessment of social, cultural, or economic factors is focused on the needs of risk analysis. The recent development of sustainability impact assessment tools (SIAT) to forecast changes as a result of policy initiatives integrates social, economic, and ecological analysis. Examples have been developed for trade sustainability impact assessment by following the two main goals of integrating sustainability into trade negotiations and informing the public and other stakeholders as early as possible. A trade sustainability impact assessment is applied on the basis of a small number of indicators by using simple calculations of qualitative and quantitative statistics. Other examples of SIAT have been developed for sustainability impact assessments of land use changes on an indicators-based frame of land use functions when using a set of models for each indicator analysis. Another recent SIAT development is based on the combination and formal linkage of different models to develop and test a multiperspective set of economic, social, and environmental indicators of the sustainability and multifunctionality of systems, policies, and innovations in agriculture and agroforestry. When SIAT are used, ERA is developed further through a general sustainability analysis. New evaluation frameworks combine participatory tools, environmental appraisal tools, multicriteria analysis, monetary assessment tools, and scenario analysis on the basis of models and physical assessment tools.

Burghard C. Meyer

See also Environmental Impact Assessment; Environmental Planning; GIS in Environmental Management; Participatory Planning; Regional Environmental Planning; Risk Analysis and Assessment

---

**Further Readings**


MULBO (Multikriterielles Landschaftsbewertungs- und Optimierungsverfahren; Multicriteria Landscape Assessment and Optimisation): www.mulbo.de


SEAMLESS (System for Environmental and Agricultural Modelling: Linking European Science and Society): www.seamless-ip.org

Soil & Water Assessment Tool: www.brc.tamus.edu/swat

E-COMMERCE AND GEOGRAPHY

Electronic commerce (e-commerce) is generally defined to encompass any commercial activity that uses the transmission of electronic data to facilitate buying and selling. E-commerce therefore spans a wide variety of electronic data to facilitate buying and selling. E-commerce therefore spans a wide variety of industries, not only those trading in tangible goods (as traditional brick-and-mortar storefronts do) but also those brokering the transfer of intangible ownership rights. It also uses a wide variety of technologies, ranging from the simple buying and selling of products using electronic mail to more complex inventory-tracking systems that automatically replenish stock based on sales. Over the past 10 years, e-commerce has become an integral part of business organization and economic strategy—accounting for more than $3.3 trillion worth of transactions in the U.S. economy in 2007 (Census 2009). E-commerce is of particular relevance to the study of geography as it enables (though it does not guarantee) change in the spatial organization of the economy. This entry reviews how e-commerce can help overcome the frictions of geographical distance while operating as a centralizing force throughout the global economy.

E-Commerce and the “Death of Geographical Distance”

As e-commerce and other information technologies became more widely available, many observers predicted the “death of geographical distance,” a concept that gripped both the popular and the
policy imagination. The core supposition is that anyone anywhere can engage in the global market. This idea is based on the concept that electronic technologies remove the geographic barriers to entry that previously prohibited some people and some communities from full participation in economic activities. Indeed, some prognosticators have proposed that certain industries will cease to conduct sales at physical locations, removing the need for large sums of capital investment before starting business. These discourses are compelling because of their egalitarian promise, but they greatly oversimplify the complex relationship between distance, electronic technologies, and the economy.

Despite the rhetoric, bank branches and retail stores still operate, and would-be entrepreneurs still face many of the same barriers to entry. There are several reasons why online exchanges have not rendered physical locations completely obsolete. First, many economic activities rely on the knowledge embodied in skilled labor, capabilities that are most easily exchanged (and regulated by managers) over minimal physical distances (at least among today’s workers). Second, Internet adoption rates vary widely throughout the world. In some countries, such as Sweden or Canada, more than 75% of the population has regular Internet access, while in other countries, such as India or Indonesia (not to mention sub-Saharan Africa), the rate is less than 10%.

Perhaps the most important reason why e-commerce and other information technologies are unlikely to eradicate geographical barriers is that moving physical products through e-commerce is not as easy as transferring data. For example, moving tangible products from one place to another imposes logistical costs that might favor centralized distribution to a large number of customers (e.g., through traditional storefronts) rather than decentralized distribution straight to each buyer. Similarly, consumers may feel more confidence when buying some products, such as clothing or appliances, if they can see and feel them before making a purchase. Although some products are available for instantaneous exchange and are sufficiently standardized for consumers to buy them remotely with confidence (e.g., music, video, or e-book downloads), these products are relatively few.

**E-Commerce as a Centralizing Force**

Although the death of distance has been overstated, e-commerce does help ameliorate the friction of distance—sometimes by directly replacing traditional sales channels but more importantly by operating in the background to make them proceed more efficiently. Many individuals engage in e-commerce, of course—either shopping directly through online retail stores or buying and selling items through intermediaries such as eBay or Amazon. These exchanges may be the most visible form of e-commerce, and they have experienced positive growth in many countries around the world, but they actually make up a rather minor portion of it. For example, retail transactions, also known as business-to-consumer (B2C) sales, account for only a quarter of e-commerce shipments in Canada and a negligible 3.2% of total retail sales in the United States.

The bulk of e-commerce activities are concentrated in the manufacturing and wholesale sectors—which is to say business-to-business (B2B) sales. E-commerce can make it easier for businesses to overcome the friction created by physical distance. It provides a cheaper means of communication with workers, suppliers, and customers located in various places around the world. For example, when a customer buys an item from Walmart, the product supplier (who may be located on the other side of the world) immediately receives notification of the point-of-sale purchase and can begin processing a new shipment for that particular store. Although this is a vast oversimplification of an extremely complex supply and distribution chain, by changing the organization and geography of their distribution network, companies such as Wal-Mart successfully blend e-commerce with the physical movement of goods.

The use of e-commerce is not limited to large multinational corporations. Small- and medium-sized enterprises around the world have also adopted e-commerce activities. In Thailand, for example, silk merchants used e-commerce to widen their distribution channels, attract new customers, and facilitate more seamless transactions with existing customers. However, the integration of e-commerce activities among smaller businesses does not tend to be as expansive as it is
Economic base analysis is a mainstay in economic geography and regional development. It stems from a simple theory of regional economic growth: that the growth of cities and regions is tied to their success in exporting goods and services beyond their borders.

Economic base analysis is a demand-driven theory that divides the regional economy into two sectors: a “basic” sector that fulfills demand originating from outside the region and a “nonbasic” sector that satisfies demand from within the region. The theory posits a strong and stable relationship between the basic and nonbasic sectors such that any change in the basic sector produces changes in the nonbasic sector sufficient to restore their proportional relationship. The relationship, or linkage, between the two sectors takes two forms. First, the nonbasic sector is assumed to supply the basic sector with many of the goods and services it needs to produce its exports. Second, workers in the basic sector demand a variety of residential goods and services (e.g., groceries, dry cleaning, utilities) that the nonbasic sector supplies. None of these residential goods and services would exist, according to the theory, in the absence of the basic sector.

The economic base model seeks to quantify the relationship between the basic ($B$) and nonbasic ($NB$) sectors that constitute the total economy ($T$). Typically expressed in terms of employment or earnings, the simplest economic base model is expressed as

$$T = B + NB.$$  \hfill (1)

Assuming a stable relationship in the basic-nonbasic ratio, nonbasic activity is proportionately related to total activity, or

$$NB = aT.$$  \hfill (2)

Substituting (2) into (1) and rearranging yields

$$T = [1/(1 - a)]B,$$  \hfill (3)

where $a$ is the propensity to spend locally. The term $[1/(1 - a)]$, equal to $T/B$, is called the economic base impact multiplier. While the basic-nonbasic relationship is theorized to be stable over time, it is not assumed to be the same across all regions.

Typically, the economic base model is used for economic impact analysis. Suppose, for example, that a region successfully attracts a new manufacturing plant that will employ 300 people and sell to national markets. Furthermore, suppose the regional impact multiplier to be 1.75. With sufficient excess capacity, the total impact on the region would be $(300 \times 1.75) = 525$ jobs, that is, 300 jobs in the basic sector and 225 jobs in the supporting nonbasic sector.

A critical step in economic base analysis involves dividing the regional economy into basic
and nonbasic sectors, a process called bifurcation. The four most commonly discussed methods for identifying exports are assignment, surveys of local businesses, the location quotient method, and the minimum requirements approach. There are also several variations and hybrid approaches.

The assignment method simply relies on the analyst’s judgment about which sectors, or portions thereof, serve external markets. It is common for federal government, agriculture, mining, and manufacturing to be assigned to the basic sector since activity in these sectors is usually exogenous to the region, especially in small or rural areas. Though there is no theoretical basis for the assignment method, a knowledgeable analyst might produce a reasonably accurate description of the region’s economic base. At the other end of the spectrum is a complete survey of businesses in the area to determine what portion of their output is sold to customers residing outside the region. Because complete surveys are costly, partial surveys are sometimes used to inform the assignment method.

Location quotients are probably the most common technique used to identify export activity. A location quotient is a measure of relative concentration. It compares the proportion of activity in each regional industry to the proportions found in a reference region, usually the nation. If a region’s proportion of employment in, say, construction is greater than the national average, the excess is assumed to be exported. Using employment as the numéraire, location quotients are calculated as

\[ LQ_{ir} = \frac{E_{ir}}{E_{rn}} \tag{4} \]

where i, r, and n refer to the industry, region, and nation, respectively. For all \( LQ_{ir} \) greater than 1, export employment, \( X_{ir} \), is determined by

\[ X_{ir} = \left[ 1 - \left( \frac{1}{LQ_{ir}} \right) \right] E_{ir}, \quad \forall LQ_{ir} > 1. \tag{5} \]

Adding all the \( X_{ir} \) terms together produces an estimate of the basic sector, or

\[ B = \sum X_{ir}. \tag{6} \]

Location quotients require several assumptions:

1. Labor productivity in the region and that in the nation are equal.
2. Regional and national consumption patterns are identical.
3. There is no cross-hauling (i.e., the region does not import and export the same commodities).
4. The nation is not a net exporter.

Of the four, the cross-hauling assumption is the most problematic. Cross-hauling is commonplace, and in its presence, the location quotient provides an estimate of net exports rather than gross exports, which is desired. In such cases, the size of the basic sector will be underestimated and the resulting impact multiplier inflated.

The minimum-requirements approach is a variation of the location quotient method but uses similar-sized regions as the basis for comparison. On an industry-by-industry basis, the analyst finds the smallest proportion of activity in each sector within the set of comparison areas. The smallest proportion is thought to represent the minimum level of production required to satisfy local demand. The remainder is assumed to be exported.

Empirically, the economic base model is sensitive to aggregation, both industrial and geographic. The most reliable results are obtained for relatively small regions using the most detailed industry data available. Furthermore, the basic-nonbasic relationship is not stable over time, and so the model should not be used for long-run projections. Perhaps the greatest shortcoming of economic base analysis is its focus on exports as the sole source of growth, ignoring the roles of consumption, investment, and government spending as sources of demand.

In spite of its shortcomings, economic base analysis remains heuristically useful for students and practitioners, offering a simple growth model that can be readily extended to incorporate additional relevant variables. Furthermore, advances in spatial data analysis and spatial econometrics may breathe new life into the simple model.

Harrison S. Campbell Jr.
ECONOMIC GEOGRAPHY

Economic geography is one of the most diverse, vibrant, and catalytic subdisciplines within human geography. It is concerned with describing and explaining the varied places and spaces in which economic activities are carried out and circulate. It was institutionalized as a subdiscipline during the late 19th century in both Western Europe and the United States. Initially connected to projects of empire (especially in its earliest form, commercial geography), economic geography has since been through a series of intellectual transformations, including a regional approach, spatial science, radical political economy, and, most recently, a “cultural turn.” Each new framework, however, has rarely eradicated the previous one. Rather, economic geography is more like a palimpsest, with previous versions and approaches to the discipline continuing to remain at least partially visible in present incarnations. The contemporary version of the discipline is now highly variegated, with research conducted on a wide front and reflecting the changing and expansive character of what has increasingly become its object of study, global capitalism.

### History

Existing in embryonic form as commercial geography, the discipline was formally defined in 1882 by the German geographer Götz. While commercial geography served immediate practical purposes, the goal of economic geography was loftier. It was to be a science, to explain how the character of different geographical areas influenced the production and movement of goods. By the 1890s, economic geography courses appeared in U.S. and British university calendars. In 1896, George Chisholm began lecturing at Birkbeck College, London, using his own economic geography text, *Handbook of Commercial Geography*, the first ever written in English. And in 1903, Lionel Lyde, at University College London, was appointed the first professor of economic geography. The discipline was up and running.

Chisholm’s book (published in 20 editions) and J. Russell Smith’s later American version, *Industrial and Commercial Geography*, set out an initial disciplinary agenda and intellectual style that continues to resonate. The original colonial context of the two books no longer applies, but remaining germane are their concerns with empirical detail and numbers, their predilection for geographical categorization made visible by the map, and their tracing of relations among places through economic flows, especially commodity flows.

Chisholm and Smith stressed the physical character of goods, particularly those of natural resources, with spatial relations or physical location coming second. The American geographer Ray Whitbeck argued it should be the other way around: The emphasis should be on the country, not the commodity. That increasingly became the case, at least in U.S. economic geography after World War I, when a regional approach took hold. Regions were portrayed as unique and obvious once the descriptive typologies used to pigeonhole
Regions remained part of the economic geographical lexicon but were conceived utterly differently—as explanatory, theoretical, and instrumental, spatial units to achieve functional objectives, not to be “merely” described. Clarence Fielden Jones and his eightfold typological scheme for classifying regions were no longer recognizable as part of the discipline.

This often abstract, closed, and narrowly conceived discipline did not last. It was out of sync with economic geography’s own history and, by the late 1960s, increasingly out of sync with its wider historical moment. Radical geography emerged during the early 1970s, propelling economic geography in a different direction. David Harvey’s classical Marxist theorizing of capitalist accumulation and crisis was important, but even more so was Doreen Massey’s work on industrial restructuring, underpinned philosophically by critical realism and culminating in the locality project that dominated British economic geography during the late 1980s. Across the Atlantic, a group of radical economic geographers in California carried out important empirical and theoretical work, not on cases of industrial capitalism’s decline (much of the focus of the locality project was a deindustrializing Britain) but on its successes, such as high-tech industries in places such as Silicon Valley and Orange County outside Los Angeles. British and American interests merged in the late 1980s, producing discussions around the transformation of an old, disintegrating Fordism, defined by assembly-line techniques and the mass production of a standardized product, into a new, emerging post-Fordism, defined by flexible workers and machines engaged in batch production of differentiated products. Heavily influenced theoretically by French regulation theory, which was more open-ended, less abstract, and deterministic than classical Marxism, the economy remained central, but it was softened at the edges by recognition of the social institutions, and even culture, in which it was embedded.

Cultural and embeddedness soon became keywords of the “cultural turn” entering economic geography from the mid 1990s. Often drawing on poststructural theory, there was an attempt to rethink both the larger nature of the discipline—its empiricist epistemology, its entrenched masculinism, its narrow methods, and its economistic
logic—as well as particular substantive topics such as labor and work, financial and business services, consumption and retailing, and the firm. Proponents argued that this turn was not simply intellectual fashion but a reflection of fundamental changes occurring within capitalism itself as it moved to a “soft,” “reflexive,” or “symbolic” form where the line between culture and economy was not just hard to see but was no longer there.

The cultural turn is far from dominant, however, and already criticisms have been mounted, often with alternatives or versions of old alternatives proposed. In truth, there is no received view, orthodoxy, or standard paradigm within economic geography. But while there is no single leading approach, or maybe because there is no leading approach, the discipline is marked by vibrancy, experimentation, and fecundity, as is evident in several areas of contemporary substantive research.

Globalization and Neoliberalism

Perhaps no research topics in contemporary economic geography are more energetically pursued than globalization and neoliberalism. The two are not necessarily connected, but over the past decade, they were often twinned. In many ways, globalization was made for economic geographical study and seen in the very first texts of the discipline, those by Chisholm and Smith. By the 1920s, however, this focus weakened. It was not until 60 years later, with the emergence of a new international division of labor, the increasing number, size, and dominance of multinational corporations, the growth of international financial capital, and new forms of communication and long-distance transportation, that economic geographers began again to deal systematically with the world as a whole. Peter Dicken, in his 1988 book *Global Shift* (2007 marked its fifth edition), was the first contemporary economic geographer to do so. One of his critical points, corroborated in subsequent economic geographical studies, is that globalization does not seamlessly eradicate geographical difference. Rather than “the end of geography,” as some economists maintain, processes of globalization show only geography’s continuing importance. Geographical differentiation is the very precondition for globalization’s possibility and achievement. In tandem with discussions of globalization has also gone a critical examination of neoliberalism, the political ideological project of promoting free markets as the optimal means for undertaking economic activity. While originating as an idea in the 1920s, it began to be practiced politically during only the 1980s. Neoliberalism’s central injunction is that places where free markets either do not exist or are insufficiently free must be transformed to meet neoliberal strictures. This means enacting policies such as trade and financial market liberalization, deregulation, privatization, and the securing of private property rights. The global consequences of neoliberal (de)regulation have been critically examined by economic geographers in studies of international finance, the internationalization of education, the international migration of skilled labor, and international trade in primary resources including diamonds, lumber, and water.

The Firm and Finance

Understanding why firms locate where they do and how companies organize economic space has long been a key theme of the subdiscipline, mostly associated with location theory. The traditional approach was based on individual case studies listing specific factors influencing firm location. But with spatial science, some economic geographers turned to complex mathematical equations. The firm was represented as an unproblematic entity working according to the laws of the market or a mathematical node in geometric space. Even during this approach, however, another school of thought developed, the geography of enterprise, concerned with treating the firm as an institution embedded in its social and cultural context. Later expanded under the cultural turn, the aim was to open up the firm to understand what goes on inside. Important themes of study include analysis of the firm’s varied production and knowledge networks (and now increasingly internationalized), the connections between firms and regional development, gender relations within the workplace, and finally the cultures of company business strategy. Connected to scholarship about the firm is also the study of finance. Finance geographies investigate the relationships between
space, place, and money, focusing on topics such as the restructuring of the financial services industry and the continued dominance of international financial centers such as New York City, London, and Tokyo. Also analyzed are offshore financial centers, such as the Cayman Islands, and their role in economic development.

### Nature and Environment

While the interests of economic geographers in nature start with Chisholm’s and Smith’s first textbooks, the contemporary interest in the topic likely dates from David Harvey and his work in the early 1970s. He was concerned to set nature within the larger context of capitalism (an approach further developed by his student Neil Smith). Current research includes studies of particular natural resources, such as water, forests, agricultural crops, and minerals, often stressing social justice and analyzing which groups of people control, access, exploit, or benefit from the transformation of nature for economic purposes. Furthermore, the recent move in the West to “go green” has spawned new trends of research around the environmental impacts of economic activity. Economic geographers are examining topics such as environmentally conscious consumers, green supply chains such as the fair-trade movement or organic farming, carbon finance or carbon taxes levied by governments, and environmental governance. Together, research agendas on both nature and the environment connect small-scale economic activity, such as the purchase of a commodity, to larger global issues such as climate change or land degradation (and link the pursuit also to political ecology).

### High-Tech and Creative Economies

Economic geographers have noted that the economic growth of the 1980s and 1990s, particularly around creative industries (such as film and fashion) and high-tech industries, resulted in new spatial dynamics on the economic landscape, especially in North America. Rather than being dominated by large corporations, these industries thrived on decentralized business systems and flexible production networks that tended to cluster in certain places, for example, Silicon Valley, outside San Francisco, and Route 128, on the outskirts of Boston. Explaining such agglomerations of high-tech and creative industries has involved partly going back to an older conceptual framework of industrial districts. Industrial districts were specific areas within a city associated with a tight-knit assemblage of small firms involved in the manufacturing of the same product. Contemporary high-tech and creative industry agglomerations follow a similar pattern. Firms within them tend to be small or medium sized and closely linked (networked) with other firms in the cluster both formally and informally. Richard Florida expands this idea in his work on creative industries by relating such clusters to particular districts within the city that are defined by a specific workforce, the young and well educated. He argues that creative industry workers,

Revolution in transportation and communication technologies since World War II have provoked the thesis of a shrinking world, or “time-space collapse.” With respect to transportation, container traffic, new, larger, and faster cargo vessels, and air cargo have been critical innovations, dramatically changing the geography of global competitive advantage. Hand in hand have also gone changes in communication modes, facilitated by computerization: e-mails, fax machines, pagers, blackberries, cell phones, and more. They have fundamentally altered business and its geography, transfiguring financial operations and flows of information and capital and permitting head offices to be half a world away from production facilities. The result is myriad new commercial products and services.

### Transportation and Communication

Transportation and communication is again a topic found in the earliest textbooks of the sub-discipline. Later, under spatial science, transportation and communication studies became effectively a branch of mathematics, topology. But even then broad questions were still asked about the relation of different forms of transportation and communication and geographical patterns of economic activity on the ground.
or the “creative class,” are drawn to edgy, diverse, gentrifying neighborhoods, which in North America primarily means the inner city. Hence, creative industries now predominately locate according to the residential patterns of the gentrifying creative class.

**Labor and Work**

Until recently, labor and work were given little attention in economic geography. While workers are a necessary input to production, they were viewed as passive and hence not very interesting. That has changed partly because one of the effects of globalization has been to mobilize labor in radically different ways. Economic geographers now study both ends of an internationally mobile labor market. On the one hand, they are concerned with (literally) high-flying professionals, such as accountants, bankers, and CEOs, who work in large multinational corporations and, on the other hand, with immigrants working in low-end jobs, as janitors, in back-of-the-house hotel positions, and in sweatshops. Both types of labor come together in world cities but in radically different spaces. A second spur to the study of labor in economic geography has come from feminism. In the late 1960s, spatial science showed that women in Western countries faced different geographical labor markets and consequences (especially in commuting). Doreen Massey’s political economy went further, demonstrating that the gendered character of labor markets was central to understanding the fundamental industrial restructuring occurring during the late 1970s and early 1980s. Furthermore, under the cultural turn, a new focus on the body as an important agent in production processes has produced a substantial literature. Its larger point is that labor and work are not neutral, merely a technical input into production, but instead are corporeally inscribed by the social, cultural, and geographical contexts in which work is practiced.

**Conclusion**

The immense range of economic geography is clearly marked by the diversity of these core areas. They capture many of the dynamic and vast shifts occurring in contemporary economic space, particularly as they connect to global capitalism. Such an open-ended approach to scholarship makes the discipline inclusive toward a variety of researchers, methodologies, and topics of study and often an exciting place to be. However, others suggest that the very diversity that gives economic geography its vibrancy is producing a subdiscipline without a central core. In some cases, the works of some scholars are incomprehensible to others in the community. The future challenge for economic geographers will be to continue to keep pace and navigate the inevitable shifts in economic space while maintaining a common ground on which to speak to and learn from each other.

*Trevor J. Barnes and Jayme Walenta*

See also Agriculture, Industrialized; Business Cycles and Geography; Class, Geography and; Clusters; Colonialism; Commodity Chains; Comparative Advantage; Consumption, Geographies of; Debt and Debt Crisis; Deindustrialization; Dependency Theory; Development Theory; Division of Labor; Ecological Economics; E-Commerce and Geography; Economic Base Analysis; Export-Led Development; Finance, Geography of; Flexible Production; Fordism; Globalization; Growth Poles; High Technology; Import Substitution Industrialization; Industrial Districts; Industrialization; Industrial Revolution; Informal Economy; Information Society; Innovation, Geography of; International Monetary Fund; Knowledge, Geography of; Labor, Geography of; Location Theory; Marxism, Geography and; Modernization Theory; Money, Geographies of; Neocolonialism; Neoliberalism; New International Division of Labor; Newly Industrializing Countries; Offshore Finance; Political Economy; Population, Environment, and Development; Ports and Maritime Trade; Postindustrial Society; Poverty; Producer Services; Regional Economic Development; Regional Science; Research and Development, Geographies of; Resource Economics; Resource Geography; Restructuring; Retail Trade, Geography of; Rural Development; Structural Adjustment; Sustainable Development; Technological Change, Geography of; Telecommunications and Geography; Transnational Corporation; Transportation Geography; Underdevelopment; Uneven Development; Urban and Regional Development; Urban Geography; Urbanization; World Bank; World-Systems Theory
Economies of scale refer to the reductions in cost that firms achieve by producing in larger rather than smaller volumes of output. “Economies” in the context refers to the benefits incurred by reducing costs, largely by spreading fixed costs over a greater quantity of output.

Mass production occurs through the standardization of parts and a detailed division of labor. Specialized divisions of labor, however, require a relatively large scale of output because a large pool of workers is generally necessary. Scale economies operate when increases in factor inputs generate disproportionately larger increases in output or, in more technical terms, the production function is not linear. For example, if a firm increases its inputs of labor and capital by 20% but sees its output rise by 30%, it enjoys economies of scale. Thus, they represent the opposite of diminishing returns in the production process.

Economists portray scale economies as a curve of long-run average costs (Figure 1), which graphs the unit costs as a function of scale of output. As unit costs decrease, they reach an optimum point and ultimately began to increase, reflecting the diseconomies of scale (diminishing marginal returns to scale) that occur when a firm becomes too large to manage and operate efficiently.

Economies of scale tend to favor the formation of larger firms and hence relatively oligopolistic market structures (those dominated by a few giant companies). Large firms generally pay much less for material inputs than do small firms because they buy in bulk and often enjoy economies of scale in transportation as well as in the production process. The presence of economies of scale varies.
widely among firms and industries. It is indisputable in sectors such as industrial agriculture and capital-intensive forms of manufacturing, such as steel and automobiles. The degree to which services, with intangible outputs, enjoy economies of scale is less clear.

Economic scale is closely intertwined with geographic location. Indeed, the choice of location cannot be considered in isolation from scale and production technique. Different scales of operation may require different locations to give access to markets of different sizes. Conversely, location itself can influence the combination of inputs and, hence, the technique adopted. Economies of scale tend to favor a select group of geographic locations over dispersed production patterns.

Barney Warf

See also Agglomeration Economies; Economic Geography; Economies of Scope; Location Theory

Further Readings

frequently found in marketing and distribution, where they underscore strategies such as product bundling.

Economies of scope also occur when there are cost savings generated by the production of by-products, when the production of one good automatically triggers the production of another. For example, a beef producer may also generate leather, and a lumber company may also create sawdust.

Potential economies of scope underlie the mergers, acquisitions, and takeovers of firms seeking to diversify into new product lines and markets. In this case, the attraction is the ability to reduce costs by operating two or more businesses under the same corporate umbrella.

Economies of scope and scale are often inversely related, particularly as firms face a “make” or “buy” decision regarding their inputs. For example, firms may subcontract (“buy”) inputs rather than make them “in-house” when they face rising uncertainty or rapid change in products or technology, when the labor process resists easy automation, or when the optimal scales of operation of production processes are markedly different. Thus, from the transactions costs perspective, externalization allows external economies of scale to replace internal economies of scope. By externalizing, firms substitute variable costs for fixed ones and spread the risks of production over their subcontractors. Achieving this is a particularly vital role of economies of scope during peak periods of demand.

Barney Warf

See also Agglomeration Economies; Economic Geography; Economies of Scale

Further Readings


There is an increasing realization that natural resources do not exist in isolation but interact with each other. Understanding this has led to a focus on a more holistic approach of managing whole ecosystems that considers both biotic and abiotic components of the ecosystems. Ecosystems are of different sizes and can be identified at multiple scales in a hierarchy. The smaller systems are embedded, or nested, within larger macroscale ecosystems. The larger systems are the environments that control their behavior. By understanding the larger forces that create macroscale ecosystems, we can better predict how management practices will affect smaller local systems. At the macroscale, ecosystem patterns are controlled by macroclimate (i.e., the climate that lies just beyond the local modifying irregularities of landform and vegetation). Over large continental areas, macroclimatic units—also termed ecoclimatic zones or ecological zones—are reflective of those ecosystems at the macroscale that biogeographers have recognized as ecosystem regions, or ecoregions. Based on macroclimatic conditions, they have developed systems that classify the natural ecoregions of Earth (Figure 1). Ecoregion classification is hierarchical in that mapping is accomplished at different scales. These classification systems are similar but differ in the hierarchical arrangement used.

The fundamental question facing ecological land mappers is “How are the boundaries of different size systems determined?” Delineating units involves analyzing the controlling factors that cause ecosystem patterns at various scales and then using significant changes in controls as boundaries. Areas of uniform climate are used to identify ecosystem units because climate acts as the primary input of energy and moisture into the system. As the climate changes, the kinds and patterns of dominant life forms of plants and animals change, as do the kinds of soils. Ecosystems of different climates differ significantly. Controls over the climatic effect change with scale. Understanding these controlling factors and the scale at which they operate is key to setting ecosystem boundaries.

Ecoregion subzones or provinces such as savanna, steppe, or tundra correspond to major
plant formations, which are delimited on the basis of features of the vegetation. These subdivisions express more refined climatic differences than do the zones. Mountainous provinces are distinguished within a zone where, as a result of altitude, the climate differs sufficiently from the adjacent lowlands to cause a typical sequence of altitudinal belts.

Within an ecoregion, landforms, through varying height and degree of ground surface inclination, further modify the macroclimate to local climate to form repeated patterns of smaller ecosystems, called sites. By observing the behavior of the different kinds of systems in a region, it is possible to predict the behavior of an unvisited one. Hence, an ecoregion’s map can be used to spatially extend data obtained from limited sample sites. The results of observations at representative sample sites from each ecoregion would be potentially useful in detecting and monitoring global change effects.

Robert G. Bailey

See also Biogeography; Climate Change; Climate Types; Ecological Mapping; Ecological Zones; Ecoshed; Ecosystems; Ecotone; Environmental Management

Further Readings


ECOSHED

An ecoshed is a geographical niche network of community food webs practicing adaptive watershed rehabilitation. The ecoshed concept is similar to a social ecosystem, except that the culturally defined geographic area is set by the watershed boundary. *Ecoshed* is a contraction of the terms *ecology* and *watershed*. Local community food webs are interconnected networks in which each ecoshed can exchange food, goods, and services. Healthy watershed rehabilitation management based on integrated river science foundations will form resilient ecosystems. Local cultural restoration results when community food webs empower its human members through shared goals. Ecoshed processes can reestablish local control and knowledge relationships in food supply systems that have become distorted by increasing distance (physical, social, and metaphorical) between producers and consumers.

An ecoshed uses bottom-up strategies to rehabilitate watersheds through community participation. Top-down government policies are often tied into capitalistic limitations and thus take many years to become established. The concept is analogous to how organic food certification was eventually established from small groups of empowered citizens growing their own organic food and communities without government subsidies.

An ecoshed establishes a boundary within the watershed of a community-based food consumption framework. Localization of food systems and chains implies that food should be consumed as close to the point of origin as possible. The term *foodshed* has been used to include all food within a certain distance. As defined by Gail Feenstra, a foodshed is an area that is defined by a structure of supply and the foods that can be grown within it, as well as the social and cultural elements of a community. An ecoshed reflects deep local watershed ecosystem knowledge and passively integrates communities with global ecological degradation. Instead of the global powers dictating the value of ecosystem services, the ecoshed concept works with local cultures and economies from the bottom up.

Ecosheds can meet the needs of its citizens through locally owned enterprises empowering communities to have greater control over their basic need of foods and services. The profits are recycled back into the community, stabilizing the economy and providing incentives to protect the environment. The full cost pricing of energy, materials, and land use will expose the true inefficiencies of conventional agriculture, making ecosheds the better alternative.

The ecoshed concept may seem abstract and faces challenges to adoption; however, many foundations are working to create proper distribution and marketing systems. Millions of enterprises and public initiatives that value local food economies at the fringes of the global economy are gaining momentum. Local grassroots food movements such as Slow Food, community-supported agriculture, and farmers’ markets are already established. A healthy watershed will evolve from a successful ecoshed objective. The basic healthy food and clean water benefits of the ecoshed are the first drops in the bucket to solving global environmental crises such as climate change and river rehabilitation. Through integrated adaptive watershed rehabilitation, local food webs, and social restoration, the synergisms to achieving ecosystem integrity will spin food webs of ecoshed districts.

*Benjamin Newton*

See also Bioregionalism; Ecoregions; Ecosystems; Food, Geography of; Organic Agriculture; Permaculture; Watershed Management

Further Readings

**Ecosystem Decay**

Ecosystem decay is a term coined by the conservation biologist Thomas Lovejoy to describe the process by which species become locally extinct and ecosystem functions deteriorate following habitat fragmentation. The term has become synonymous with an ongoing, ambitious study of ecosystem decay in the Amazon rain forest, initiated by Lovejoy and colleagues in the 1980s. In this study, forest fragmentation has been found to have unexpected and cascading impacts on species remaining in forest fragments. This entry first describes the process of ecosystem decay and then examines impacts on the Amazon rain forest that have been identified by researchers with the Biological Dynamics of Forest Fragments Project.

It has been long recognized that habitat fragmentation is a leading cause of species loss worldwide. This widespread threat is also referred to as “islandization,” whereby a fragment of untouched habitat is isolated in a sea of altered habitat. Nature reserves have been ongoing testament to the negative effects of islandization; for example, populations of mammals with large home ranges visibly decrease within reserves, and this decrease is relative to the size of the reserve. Nonetheless, the mechanisms by which most species cease to exist in such habitat fragments have been difficult to isolate within the complexities of an ecosystem, a difficulty recognized by Lovejoy when he labeled this complex process as ecosystem decay.

To grasp the impacts of habitat fragmentation, biogeographers turned to island biogeographic theory soon after it emerged in the 1960s. By understanding the impacts of area, isolation, species immigration, and extinction within virtual islands, conservation biogeographers aimed to improve nature reserve selection. However, debates over optimal sizes of reserves ensued, a well-known one being the SLOSS (single large reserve or several small) debate. The questions driving such debates were based on the unknown complexities of ecosystem decay.

Today, the term ecosystem decay is often replaced by habitat degradation or habitat loss, reflecting the tendency of biogeographers and conservation scientists to focus on habitats rather than ecosystems. Because we are still far from understanding the inner workings of ecosystems and because biodiversity loss is occurring at an unprecedented rate, it is generally more manageable to work toward the preservation of habitats when making conservation decisions. Habitats can be defined at many scales and do not necessarily require in-depth knowledge of entire systems.

**Ecosystem Decay in the Amazon Forest**

The ecosystem decay study in the Amazon rain forest is called the Biological Dynamics of Forest Fragments Project. Initially, the project targeted questions around optimal reserve size, through an in-depth, long-term study of ecosystem decay. The study began with the experimental clearing of large tracts of pristine rain forest in the early 1980s—an approach that sparked controversy at the time. The purpose of this clearing was to create a matrix of forest fragments representing different areas and different levels of isolation from intact forest. The areas of the fragments ranged from 1 to 100 ha (hectares), and the isolation ranged from 80 to 650 m (meters) from intact forest or other fragments.

Since creating the fragments, numerous scientists from around the world have monitored, and continue to monitor, the ecosystem decay occurring within these fragments. Some of the more consequential findings have to do with conditions on the edge of fragments (called the edge effect), total area and the species-area relationship, and fragment isolation.

**Edge Effects**

The conditions occurring on the edge of a forest differ considerably from those within the interior of a forest. For example, temperature alone will change dramatically as one moves from the edge of a tropical forest to its interior. If a forest is carved up into smaller pieces, then the ratio of edge to total area will be altered; the portion of
the forest experiencing edge conditions will increase as the portion experiencing interior conditions will decrease (Figure 1). The effects of such increased exposure to edge conditions (the edge effect) has been a key focus of the Biological Dynamics of Forest Fragments Project.

What Thomas Lovejoy and his team found is that many species occurring within the interior of the forest are ill adapted to the conditions found on the edge, such as increased temperature and light. While variables such as temperature and light are predictable drivers of plant species occurrence, the team found that trees responded to edge effects with an unexpected and dramatic increase in mortality rates. Additionally, understory plants, leaf litter, soil nutrients and moisture, humidity, and wind penetration were all subsequently and simultaneously altered. These effects were generally found up to 60 m into the interior.

Edge effects on animal species were equally surprising. Light exposure detracted some birds to the point where they would not emerge into the less dense edge vegetation, even if it was a means to accessing greater food resources. The many species relying on leaf litter disappeared from the edge transition zone, as did most species associated with the trees and understory vegetation that died along the edge. This means that the impacts on microfauna were more far-reaching than predicted, contradicting previous beliefs that habitat fragmentation affects large-bodied animals more than it does small-bodied animals.

In addition to the effects of edge conditions, fragments were exposed to intrusion by species from the peripheral habitat; in some cases, this included cattle grazing land, and in others, it included regenerating forest. Non–rain forest species rapidly filled the niches within the edge.
transition zone, as was expected, but in some instances, they altered species compositions as far as 200 to 400 m into the interior of the forest, as was the case with butterflies. Competition for resources, altered microhabitats (canopy structures and ground cover), and altered predation were just some of the resulting impacts on highly specialized species within the forest interior.

The negative effects of invasive species on true islands are usually the result of passively dispersed species (i.e., rats on boats), since there is no bordering terrestrial habitat from which threatening active dispersers can migrate. Hence, the findings of the Biological Dynamics of Forest Fragments Project have demonstrated some of the unique threats to habitat fragments on land.

Edge effects in habitat fragments are now widely accepted as key drivers of ecosystem decay and have been recorded in numerous habitat types. There is always a transitional zone on the periphery of a habitat fragment that is unsuitable to certain species of the interior. The extent to which ecosystem decay occurs depends on the individualistic responses of each species.

Due to the long-existing, self-regulated climatic conditions of rain forests, many species have evolved to fill highly specialized niches. Hence, the number of species affected by edge effects might be profound in a rain forest fragment surrounded by cattle grazing but less profound in a scrubland fragment surrounded by cattle grazing. So the concept of edge effect is transferable to other habitats only as a general guideline for identifying ecosystem decay processes.

**Figure 1** Edge-to-area ratios for contiguous versus fragmented areas

*Source: Author.*
greater species diversity and that therefore the larger the nature reserve, the better it was. This belief reflected a key principle of island biogeographic theory, which is that there is a negative relationship between area and extinction rates—the larger the island, the lower the extinction rate, and vice versa. Some, however, believed that species diversity depended heavily on the diversity of habitat within the fragment and not necessarily on size. While several studies set out to prove and disprove these hypotheses, the ecosystem decay study in the Amazon played an important role in supporting island biogeographic theory and the species-area relationship.

The mechanisms by which extinction rates increased with decreased area were found to be grounded in resource limitation and reproduction. Species with large range requirements, often due to the distribution of food sources, decreased much more rapidly in the small forest fragments (1 ha) than in the larger fragments (10 and 100 ha). Studies in other environments have shown how wide-ranging large mammals are often the first to become extinct in habitat fragments. The Amazon study demonstrated that large mammals as well as many small animals, such as ant-following birds and dung beetles, suffered the same fate. This extent of area-related ecosystem decay was unexpected.

The reproductive success of a species is also a function of area, particularly in species with low densities. The smaller the fragment, the fewer individuals there will be and the greater chance of a genetic bottleneck. This effect is also controlled by an area’s level of isolation. In the Amazon, many low-density species managed to move from forest fragments to intact forest—this movement was triggered first in the small fragments but eventually in all fragments. But in reserves that are more isolated (i.e., surrounded by human settlements), such movement of large mammals is often impossible, and the species become locally extinct.

**Isolation**

The extent to which a fragment is isolated from source populations of species will dictate the number of species occurring in that fragment. The further a species must travel, the less likely it is to successfully immigrate to the fragment or to leave the fragment and find greater resources. With little incoming or outgoing movement, extinctions can often outnumber immigrations, and the resulting decrease in species number causes ecosystem decay.

The significance of the immigration-to-extinction ratio is highlighted in island biogeographic theory, and isolation plays a key role in defining this ratio. Levels of isolation are relatively easy to measure for islands because it can be assumed that ocean water is a constant barrier to most terrestrial species (birds excluded). Hence, if access to an island involves the crossing of 100 km (kilometers) of sea, one can assume that this obstacle of distance will be unchanged when estimating immigration rates. However, if one tries to measure the isolation of a habitat fragment, then the continuous changes of peripheral habitats must be accounted for (i.e., occurrence of seasonal corridors, fires, flooding).

Lovejoy and his colleagues designed their Amazon experiment so that each fragment would be surrounded by cattle farming and have different distances between fragments and intact forest. In some cases, cattle grazing persisted around the fragment perimeters, but in others, it was replaced with regenerating forest. Their studies have shown that fragments surrounded by regenerating forest had significantly greater movement to and from the fragment when compared with fragments with cattle on their perimeter. The results illustrate the significance of perimeter land uses in aggravating or abating ecosystem decay and the need to identify each fragment’s ever-changing barriers instead of focusing only on distances to source populations.

The study also highlights the need to recognize species’ individual responses to varying levels of isolation in order to understand ecosystem decay. Small animals, such as certain crawling insects, were unable to cross rain forest clearings of only 15 m. Large carnivores, such as jaguars, can easily cross the distances between fragments, but human activity, especially hunting, can restrict their movement. Seemingly mobile species, such
as birds, were surprisingly constrained even in fragments bordered by regenerating forest. While ant colonies were able to move to nearby fragments, some ant-eating birds were not. Each species’ response affects another species, which affects another species and so on, a complex process of ecosystem decay that we are yet to understand fully.

Jennifer S. Lalley

See also Deforestation; Ecosystems; Extinctions; Forest Degradation; Forest Fragmentation; Island Biogeography; Land Degradation; Single Large or Several Small (SLOSS) Debate; Species-Area Relationship

Further Readings


Ecosystems: From Passive Support to Active Component

Many studies made in the fields of biology and ecology during the past few decades have progressively shown the inadequacy of a traditional representation of nature as a passive entity containing all the resources extracted by human activity and receiving all the emissions and waste produced by mankind. A description in terms of ecosystems that actively produce and allocate natural services is far more satisfactory because the role of natural systems is anything but passive. The network of ecosystems provides a great number of natural services, which are vital for the survival of the human and of the other living species. The Millennium Ecosystem Assessment (MA), following several other studies, describes the following four categories of ecosystem functions, further detailed in several ecosystem services:

1. Supporting services such as
   • the capture of solar energy and its embodiment as biomass,
   • the decomposition and recycling of organic waste,
   • the formation of soil, and
   • the fixation of nitrogen in soils

2. Provisioning services such as
   • the production of food (which makes the life of all heterotrophic organisms, including human beings, possible) and
   • the provision of biotic materials or fuel

3. Regulating services such as
   • regulation of the composition of the atmosphere,
   • climate regulation (including the redistribution of humidity),
   • regulation of the water cycle,
   • water purification,
   • pollination,
   • biological control of disease and pests, and
   • land stabilization and the control of soil erosion

4. Cultural services, including aesthetic, spiritual, educational, and recreational ones
All the services listed here contribute directly or indirectly to making our planet a suitable place to live in, for humankind and for all the other living species.

Each of these services derives from living organisms and in particular from some species that are critical in maintaining habitat stability and in supporting ecosystem services. Thus, a reduction of these species at a local level will have negative consequences on ecosystem services also at the global level.

**Resilience, Resistance, Stability, and Persistence of Ecosystems**

There are some properties that focus on the way that ecosystems respond to disturbance. Following C. S. Holling, we define resilience as the capacity of an ecosystem to return to the initial state after a perturbation, maintaining its essential characteristics. This emergent property of ecosystems refers to the ability of an ecosystem to cope with environmental disturbances (climate change, pollution, floods, fires, etc.), without changing its state into a qualitatively different level. For a high level of resilience, an ecosystem is characterized by a great capacity to survive shocks and, if damaged, to restore itself. This property is conferred at multiple scales by genes, species, functional groups of species, and processes interacting within the system. This process of restoration after disturbance helps ecosystem renewal and further development.

Resistance is defined as the capacity of the ecosystem to absorb disturbances and remain largely unchanged. If an ecosystem is resistant, it shows little alterations to withstand non-catastrophic perturbations. Ecosystems may be highly resilient but have low resistance to a given perturbation: For example, grasslands have a low level of resistance but a higher level of resilience to fire.

Resistance is related to the concept of stability, which reflects the capacity of an ecosystem to maintain a dynamic equilibrium over time while resisting change to a different state. A stable ecosystem persists when it is able to absorb disturbances and remain largely unchanged over long periods of time.

**Ecosystems and Spatial Scale**

It is important to describe the properties of ecosystems with respect to spatial scale. Many recent geographical studies have highlighted, from different points of view, the complexity and the importance of scale in contemporary debates due to the role of scale in most of socioeconomics and environmental dynamics, as described in several geographical theories and topics.

An analysis of the spatial scale of ecosystems shows that they behave in ways that can be described on a local scale, but they also have characteristics that can be understood at other levels and cannot be described in terms of local ecosystems alone. According to a deeper analysis, at least three different levels characterizing ecosystems can be distinguished.

1. **The local level:** First, ecosystems are spatially localized systems, which, although local from the territorial point of view, have permeable and blurred borders. Hence, only some of the properties that characterize these systems can be explained correctly by a local description. The actions of human origin that lead to local environmental externalities (pollution of local soil or lakes, fires in local areas, etc.) fall within this level of representation.

2. **The intermediate level (mesoscale):** This scale is taken into consideration in analyses by geographers referring to the concept of “ecoregion,” introduced by Robert Bailey in 1976 and progressively deepened by him and his collaborators. These studies propose a description not limited to the consideration of local ecosystems because such a description is unable to capture the quality and processes (orography, hydrography, climate, flora, fauna, etc.) that emerge only at the regional level. At this scale, the network of interactions and feedbacks linking the ecosystems integrates them to form an ecoregion, a natural environment that, in spite of its local variations, is characteristic of that territory (i.e., the desert zone of the Colorado Plateau, the Appalachian mountains, the Amazon forest, etc.).

3. **The global level:** As they have permeable borders, the various ecosystems exchange energy
and material through a network of relations that takes in the whole planet. The case of emissions into the atmosphere is typical: Even if they are localized in a small area, winds and atmospheric currents disperse the pollutants, which generate fallout of acid rain, ozone depletion, and the greenhouse effect, at a global level. The only level at which the system becomes really closed is the planetary level. Most of the functions, relations, and feedbacks existing among ecosystems are closed only on the global scale. It is precisely at this level that there emerge effects that do not derive from the mere addition of local behaviors but represent the result of out-and-out synergies. There are many examples of this level, including the regulation of carbon dioxide \((\text{CO}_2)\) in the atmosphere, where the \(\text{CO}_2\) emitted (locally) is not reabsorbed by the local ecosystems (except for a very small portion) but becomes part of the atmosphere, whose \(\text{CO}_2\) composition is controlled by the balance that exists only on the global scale, between the activities of photosynthesizers and respirators.

**Strategies for Better Management of Ecosystems**

From 2001 to 2005, the MA was conducted, which was called for by the United Nations to assess the consequences of ecosystem change for human well-being and to establish the scientific basis for actions needed to enhance the conservation and sustainable use of ecosystems and their contribution to human well-being. Their results offer a state-of-the-art scientific appraisal of the conditions and trends in the world’s ecosystems and the services they provide and the options to restore, conserve, or enhance the sustainable use of ecosystems. According to MA conclusions, human actions are depleting Earth’s natural capital, damaging the ability of the planet’s ecosystems to sustain future generations. At the same time, the assessment shows that, with appropriate actions, it is possible to reverse the degradation of many ecosystem services over the next 50 years, but the changes in policy and practice required are substantial and not currently underway.

Many of the different ways of improving management of ecosystems and several methodologies can be integrated with each other. One of the most important strategies is the ecosystem approach, which provides a framework that can be implemented to reach the aims of the Convention on Biological Diversity, the international treaty that focuses on ecosystems and on their problems, and in particular its three main objectives: (1) conservation, (2) sustainable use, and (3) fair and equitable sharing of the benefits of biological diversity. Thus, the ecosystem approach is a strategy for the integrated management of land, water, and living resources that promotes conservation and sustainable use in an equitable way. It is based on the application of appropriate scientific methodologies focused on the levels of biological organization, which encompass the essential structure, processes, functions, and interactions among organisms and their environment. This approach recognizes the relationships between healthy and resilient ecosystems, biodiversity conservation, and human well-being and sets out a series of principles for decision making and action regarding the different dimensions of sustainability, environmental and socioeconomic. It also recognizes that ecosystem processes are often nonlinear and that management must be adaptive to be able to respond to such uncertainties. It can be applied at any scale from the local, to the meso, to the global scale. To reach this aim it proposes new types of collaboration between civil society, the private sector, and governments.

*Marco Bagliani and Antonella Pietta*

See also Adaptive Radiation; Agroecology; Agroforestry; Animal Geographies; Biodiversity; Biogeochemical Cycles; Biogeography; Bioregionalism; Biosphere Reserves; Biota and Climate; Biota and Soils; Biota and Topography; Biota Migration and Dispersal; Coupled Human and Animal Systems; Cultural Ecology; Cyborg Ecologies; Ecological Economics; Ecological Regimes; Ecological Zones; Ecoregions; Ecoshed; Ecosystem Decay; Ecotone; Environmental Management; Environmental Protection; Environmental Services; Gaia Theory; Landscape Ecology; Resilience
An ecotone is a boundary or transition zone between adjacent ecological communities or systems. Ecotones are produced where a combination of environmental conditions and biotic interactions exceed the physiological limits of the species on either side. They may result from steep environmental gradients such as abrupt changes in temperature or soil moisture or from nonlinear responses of species to more gradual changes in environmental factors. They are a focus of current research in physical geography due to their scale-dependent nature, their effect on species population dynamics and the movement of matter and energy, their relevance to the study of ecological resilience, and their role in understanding the dynamics of ecosystems.

complexity and self-organized ecosystems, and their potential sensitivity to climate change. Recent advances in spatial technologies and techniques have allowed biogeographers and ecologists to understand better the important role that ecotones play in the pattern and functioning of the biosphere.

The structural characteristics and dynamics of ecotones are scale dependent; they may appear sharp or gradual with varying spatial scales and static or ephemeral with varying temporal scales (see images). The ecotone concept, however, is scale independent; ecotones can be identified and have been studied at various scales, ranging from edge-effect investigations of wood lots in agricultural fields, to the boundary between trees and tundra on a mountain slope, to continental-scale biome transitions, such as those between deciduous forest and grassland in North America.

Frederic Clements coined the term ecotone in 1907 from the Greek words oikos (home) and tone (tension), accentuating the interaction between the adjacent communities. However, the work of Clements and other early ecologists emphasized the importance of and interaction among species within communities, which directed attention away from the boundaries between them. Additionally, analytical methods in ecology were designed to deal with homogeneous communities, which led ecologists to further ignore transitional areas. Research that did focus on boundaries emphasized an “edge effect,” which included both positive (e.g., increases in species diversity) and
negative (e.g., increases in bird nest predation) impacts on wildlife and plant communities. The emphasis of these studies was on the effects of edges as static elements.

Within the past 25 years, four factors have shifted interest back to ecotones, emphasizing their role as a dynamic component of pattern in the biosphere. First, the emergence of landscape ecology led to the explicit consideration of interactions between vegetation patterns and environmental processes and to a focus on the effects of scale on those interactions. Patterns arise from heterogeneity, itself a consequence of boundaries between relatively homogeneous patches of vegetation, land, and water, and researchers have studied how the patterns created by ecotones affect the ability of species, matter, and energy to flow across a landscape. New conceptualizations of scale produced greater interest in the scale-dependent properties of ecotones, including the recognition that different factors constrain ecotone patterns and dynamics at different spatial scales and that the constraints are more numerous and complex at fine scales than at broad scales.

Second, complexity theory, another concept with strong ties to landscape ecology, suggested that some ecosystems and landscape patterns are self-organizing, that is, the pattern is an emergent property of positive feedback between vegetation and environment within the system. Where these properties are present, abrupt ecotones can develop over gradual environmental gradients, complex ecotone patterns can develop over a homogeneous abiotic template, and vegetation change at ecotones can be amplified or diminished. Computer simulation models have been at the forefront of this area, emphasizing the creation of complex spatial structures through self-organization, and field observations have confirmed the important role of endogenous biotic interactions in ecotone pattern and dynamics.

Third, increasing concerns over global climate change, coupled with the recognition that ecotones represent the physiological limits of species’ distributions, led to interest in their potential as indicators of climate change. An ecotone that is controlled primarily by temperature, such as some alpine tree lines, would be expected to shift upslope with warmer temperatures. A low-elevation, forest-grassland ecotone that is controlled primarily by moisture would be expected to shift downslope with an increase in rainfall. However, the usefulness of ecotones in monitoring climate change has been questioned due to their typically slow movement and to complex interactions among climatic variables that might constrain ecotone dynamics.

A final factor in the recent interest in ecotones is the development of increased computing
A meadow-forest ecotone from the Cranberry Glades Botanical Area, West Virginia
Source: L. Resler.

A view of the Grand Canyon, Arizona, illustrating two ecotones: water-riparian and riparian-desert scrub
Source: D. Butler.
power, advances in spatial technologies such as geographic information systems (GIS) and remote sensing, and improved edge detection techniques. These advancements have enhanced our ability to identify ecotones, explore and map their properties over various scales, and monitor their dynamics. Once ecotones have been identified, remote sensing provides images of the vegetation and terrain with increasingly finer resolution, allowing the assessment of multiple variables influencing ecotones. Spatially explicit computer simulation models allow an exploration of the relative influence abiotic and biotic factors on ecotone pattern and dynamics. Finally, GIS provides a platform for the coupling of models and remote sensing data, the analysis of spatial pattern, and the creation of maps to display ecotone properties and dynamics.

Matthew F. Bekker

See also Climate Change; Complexity Theory; Ecological Zones; Ecoregions; Ecoshed; Ecosystems; Landscape Ecology

Further Readings


Ecotourism

Although it may seem conceptually simple, ecotourism has suffered from the lack of a precise, widely accepted definition. The International Ecotourism Society has attempted to minimize this confusion with a short and simple definition of ecotourism as “responsible travel to natural areas that conserves the environment and improves the well-being of local people.” Many ecotourism scholars have instead argued that it is more important to focus on essential criteria to be met rather than trying to craft a universal definition. Most scholars agree that to be considered ecotourism, nature-based travel must meet at least four criteria:

1. Minimization of environmental impacts: through sensitive “green design” using locally available materials; energy conservation and use of renewable energy; reduction, recycling, and proper disposal of wastes; and regulation of visitor numbers and activity
2. **Generation of funds for conservation:** through visitor fees and, if feasible, voluntary contributions

3. **Benefits to local communities:** through park revenue sharing as well as through employment and provision of local goods and services and community empowerment through ownership or management of operations

4. **Education of visitors:** through orientation programs and interpretation by trained guides on both conservation issues and cultural norms

Unlike other traditional forms of nature-based tourism, ecotourism has been explicitly conceived as an ethical form of travel and can therefore be regarded as a subset of nature tourism that meets certain criteria. Familiar forms of nature-based tourism such as wildlife tourism and adventure tourism may or may not be considered ecotourism depending on whether they meet these criteria. The relationship between different forms of nature-based tourism is depicted in Figure 1. While nature-based tourism, like other forms of tourism, has a long history, the concept of ecotourism is relatively new. The concept arose out of a convergence of trends in the 1970s and 1980s. Researchers and scholars noted with alarm the enormous environmental, social, and economic problems created by conventional mass tourism; the rapid destruction of pristine natural areas in the developing world; and the opposition by local communities to nature conservation efforts through Western-style national parks, leading to calls for an alternative form of tourism that is more ecologically and socially responsible. This view of ecotourism blended nicely with the concept of “sustainable development,” which has rapidly gained in popularity since its introduction by the World Commission on Environment and Development in 1987.

### The Practice of Ecotourism

Ecotourism is by far the fastest growing segment of the global tourism industry, itself one of the largest enterprises of the world. From a handful in the 1980s, ecotour operators have proliferated rapidly, with virtually every major travel company offering some sort of ecotravel and millions of people taking part in eco-vacations each year. These statistics are somewhat inflated because they lump together most nature and adventure travel with ecotourism, but they do demonstrate a growing public interest in this sort of travel.

As the popularity of ecotourism has grown, so has its attractiveness to governments, particularly those in developing countries. Many governments have viewed ecotourism as less problematic than other avenues of economic growth, such as mining, logging, and ranching, and perhaps even more profitable. By the early 1990s, most developing countries were promoting ecotourism as part of their development strategy. In several countries, nature-based tourism mushroomed into the largest foreign exchange earner, regardless of whether the entire country, such as Costa Rica, or a part of it, such as the Galapagos Islands in Ecuador, was being promoted as an ecotourism destination. Even countries formerly suspicious of international contact, such as China and Cuba, have now enthusiastically joined the ecotourism bandwagon. Multilateral institutions and aid agencies such as

---

**Figure 1** The relationship among different forms of nature tourism

the World Bank, United Nations (UN), and U.S. Agency for International Development (USAID), which had always been supportive of conventional tourism as a development strategy, have now become important promoters of ecotourism as a vehicle for sustainable development. Indeed, the UN proclaimed 2002 as the International Year of Ecotourism, thereby promoting its status further. Even major conservation groups such as World Wide Fund for Nature (WWF), the Nature Conservancy, and Conservation International have played a significant role in promoting ecotourism as a win-win solution for conservation and economic development.

However, when poorly planned, unregulated, and overhyped, ecotourism, like mass tourism, can bring only marginal benefits and serious social and environmental consequences. Once, unique places such as Nepal, Kenya, and the Galapagos Islands were visited by a handful of adventurous travelers, but by the 1990s the explosion of tourists had caused large-scale deforestation in Nepal, led to detrimental changes in cheetah behavior in Kenya, and threatened the survival of exotic species in the Galapagos. As ecotourism has grown in popularity, the travel industry has seized the opportunity and in the process has largely succeeded in diluting its meaning, taking advantage of the confusion over its precise definition, through the use of heavy marketing, catchy slogans, and misleading logos and awards. With the mainstreaming of ecotourism, “softer” ecotourists have demanded more mechanized transport, easy accessibility, and high levels of services and have become less interested and concerned about conservation. Many scholars agree that much of what goes on under the name of ecotourism today is “green-washing,” that is, conventional nature tourism with cosmetic changes in marketing. Another significant and growing segment is what is called “ecotourism lite,” which may incorporate solar panels, water conservation, and waste recycling but does little to benefit or empower local communities, aid conservation, or educate visitors. Genuine ecotourism that meets all of the minimum criteria is very rare. Ecotourism practice today can be represented as a continuum (Figure 2).

Where genuine ecotourism is to be found, it is most often the outcome of conservation or community groups and dedicated and passionate individuals, and such scattered successes hold out significant promise. These successful models are usually small-scale and community oriented, with visionary individuals, conservation groups, or government agencies playing important facilitating roles in organization development, guide training, and marketing. The CAMPFIRE program in Zimbabwe has instituted revenue sharing from game reserves with local communities, building schools, clinics, and water supplies, which has helped turn former poachers into conservationists. Similar programs have helped the Maasai in some of Kenya’s national parks, who have also started their own community-run ecotours. The Annapurna Conservation Area Project in Nepal has helped form a wide network of village-level committees that run lodges, maintain

Figure 2  The ecotourism continuum
trails, and guide tourists in a self-sufficient model. Advocacy and watchdog groups, such as Tourism Concern, Tourism Watch, the Ecumenical Coalition on Third World Tourism, and the Indigenous Tourism Network, have worked with local communities to advance the cause of sustainable and equitable tourism.

**Criticisms**

Ecotourism has many critics. The idea that individual tourists, by virtue of their choice of a “green” trip, will bring about social and environmental benefits, seems to fit perfectly with the dominant neoliberal paradigm, which advocates free markets and a diminished role of government regulation. As investment becomes more deregulated, local ecotourism enterprises in developing countries cannot compete with strong foreign companies, and their viability is undermined. The dominance of external players who repatriate profits leaves little benefit for local communities. Under increasing pressure, tourism facilities in many protected areas have been privatized, and private tour operators facing competitive pressure are more likely to sacrifice sustainability and equity goals. Privatization also decreases revenue for governments, making it even more difficult to maintain well-managed protected areas. The neoliberal paradigm increasingly reinforces the view that conservation must demonstrate its financial viability, forcing governments or protected area managers to allow more and more tourists, thereby undermining long-term sustainability. Overreliance on ecotourism forces destinations to engage in destructive competition with other regions and countries, where a win for one may be a loss for another. It also puts local economies at the mercy of global market forces, and events beyond their control, such as a recession in Japan or a terrorist attack in the United States, can suddenly dry up the flow of tourists.

Ecotourism usually depends on marketing of the destination as “exotic” and “unspoiled,” often referring to locals as “happy” and “friendly” people. Such characterizations can gloss over uncomfortable realities such as poverty, crime, corruption, human rights abuses, and environmental threats that may hinder tourism prospects. This helps perpetuate fantastic perceptions of the host country or region, precluding opportunities for an improved understanding of problems and potential progress toward solutions. Insensitive ecotourists often leave behind profound sociological and cultural transformations in their wake, particularly in remote indigenous communities with little prior exposure to the “modern” world. The rush to enjoy the exotic leads to trivialization of the local culture, the sudden introduction of a cash economy, the craving for modern gadgets, and the consequent breakdown of cultural norms, particularly among the youth. Promotion of ecotourism by governments or corporations also frequently sets the stage for conflicts over traditional resource use, such as when water is diverted to the tourist resort, which further worsens with the subsequent intrusion of other exploitative forms of “development.” Indeed, concerns such as these have led an international coalition of indigenous groups to condemn the 2002 UN International Year of Ecotourism.

**Ecotourism Certification**

Ecotourism certification involves setting criteria for the social and environmental impacts of ecotourism, auditing, and awarding ecolabels to enterprises with outstanding performances. Proponents of ecotourism certification argue that good ecolabels will help the public distinguish between responsible ventures and green-washing claims and thus mobilize consumer purchasing power for more desirable forms of ecotourism. Although logical in theory, several factors hinder the effectiveness of ecocertification programs. Critics argue that since genuine or hard ecotourists are a very small niche, and the majority of the traveling public desires only soft ecotourism or ecotourism lite, the consumer demand for genuine ecotourism is too small to bring about any appreciable change in the tourism market in favor of sustainable and responsible practices. Moreover, with the proliferation of certification schemes, there is confusion about which one to use or trust. Some such as “Green Globe” (heavily promoted by the World Travel and Tourism Council, an industry group) are misleading since they merely require a declaration of intent for certification without any performance criteria.
Although there are some well-regarded certification schemes (e.g., the Certification of Sustainable Tourism program in Costa Rica and the Nature and Ecotourism Accreditation Program in Australia), standards vary widely, and a need has been felt for developing a global ecotourism accreditation mechanism. Respected conservation groups such as the Rainforest Alliance have been involved since the early 2000s in establishing the Sustainable Tourism Stewardship Council (modeled after the successful Forest Stewardship Council), but its future success remains to be seen. One important limitation of certification schemes is that they tend to focus on easily measurable “green design” criteria while doing little to address equity issues such as community empowerment. In fact, certification schemes can even work against small, community-operated enterprises since they are expensive to implement. Therefore, voluntary ecotourism certification can, at best, play a minor role in moving ecotourism toward sustainability unless reinforced by strong legislative and community empowerment initiatives.

### The Future of Ecotourism

As the concept of ecotourism has seemingly become hopelessly diluted, some experts have pronounced it dead. Others claim that ecotourism is still in its infancy and can mature and grow beyond its small niche to transform many aspects of tourism practice. Whether the potential of ecotourism can be reached depends on many factors, such as if governments adopt standards for responsible ecotourism. A deep understanding of the concept indicates that ecotourism alone is not a panacea for myriad conservation and development challenges. Mass ecotourism is an oxymoron; not only does increasing visitor numbers put more pressure on the local ecosystem, but an ever-increasing number of people engaged in long-distance travel in the name of ecotourism is self-defeating. Ecotourism ventures have commonly tried to limit visitor numbers by promoting expensive luxury vacations. But catering to the high-end market does not guarantee responsible visitors, and it increases ecological impacts and discriminates against passionate conservationists of ordinary means. An alternative way to limit visitor numbers and attract responsible ecotourists is through limiting luxury and amenities instead. Under the right circumstances, responsible ecotourism can at best provide limited economic and conservation benefits. Therefore, ecotourism planning would ideally be conducted as part of a broader strategy of sustainable livelihoods development in which local communities can make informed and empowered decisions.

Perhaps the most important step toward fulfilling ecotourism’s promise is to build a more conscious traveling public. On the positive side, there have been numerous instances where dedicated tourists have fallen in love with the destination and its people and devoted their lives to these causes, whether human rights in Tibet, rain forests in South America, poverty reduction in Africa, or infrastructure in Nepal. Tourism has thus promoted “people-to-people” contacts that have aided the cause of sustainability and justice. Among the overwhelming negative trends, there is also the positive trend of a steady increase in volunteer tourism, in which groups of volunteers work on a variety of sustainable development projects while on “vacation.” If such linkages can be strengthened, ecotourism can become a force of positive change.

Abhijit Banerjee

See also Biosphere Reserves; Conservation; Environmental Impacts of Tourism; Neoliberal Environmental Policy; Parks and Reserves; Sustainable Development; Tourism

### Further Readings

Studies of geographies of education explore the complex interactions between education, space, and civil society. The topic is distinct from geographic education, which deals primarily with the teaching and learning of geography as a subject. In recent years a vibrant resurgence in this field has resulted in a growing body of work, theoretically and empirically dispersed within the discipline itself. Critical studies within this broader realm are concerned primarily with social justice, power relations, and structural inequities relating to or emanating from the educational system and have drawn on Marxist, feminist, and poststructuralist approaches to explore wide-ranging topics.

Educational spaces are imbued with multiple purposes and meanings, with the ideologies of the state and its corresponding discursive and material realities becoming discernible. As a lifelong learning process, education is an ongoing activity tied to the production and consumption of knowledge. Including early childhood education, primary to secondary schooling, postsecondary institutions, formal to informal learning, and collective to individualized prospects, the geographies of education reflect various sites and scales of opportunities to investigate an array of geographical concerns. As democratic public spheres, public schools are sites for justice, equality, freedom, and social transformation. Ronald Manzer defines schools as human communities, public instruments, and political symbols and the means by which people in a political democracy collectively strive for civic virtue, economic wealth, and cultural survival. Apart from its educational mandate, these are places where neighborhood integration, social capital formation, and the fostering of civil society are ideally endorsed. The city-school relationship is also intrinsically linked to the planning and sustainability of urban regions through the quality and vibrancy of its educational institutions. Cities provide the context for communities of difference and have brought educational institutions to the forefront of these debates as sites of empowerment and social cohesion. In the smaller towns, the threat of declining populations and out-migration due to the closure of manufacturing plants and resource-based industries has had detrimental effects on the educational prospects of these places. In rural areas, schools provide one of the few opportunities for community engagement and civic participation, and the threat of school closures has widespread and long-term implications on the well-being of its citizenry.

Critical scholars have long argued, however, that in capitalist societies, educational spaces are sites of cultural and social reproduction succumbing to dominant class value systems. Relegating the educational sphere to the logics of the market and the pursuit of profit—for instance, the call for privatization and choice and the streamlining of the production of knowledge into professional, vocational, and technically trained workers—signals the manipulation of workers into complacency and servitude to capitalist regimes. These forms of institutional discrimination practices are reflected in the spatial segregation and redlining of schools, particularly along racial, class, gender, and disability divides. Similarly, the system of education that was imposed on indigenous peoples by colonial and neocolonial powers led to the historical and spatial processes of their marginalization, dependency, and humiliation. These exclusionary practices in turn have led to spaces of stigmatization and subversion and to the demoralizing and silencing of many marginalized communities. Spaces of education, however, are also sites of contestation, resistance, and possibility. As Ronald Manzer notes, public schools are not only the objects of domination and the products of compromise,
they are also potentially agencies for creating political consensus. Thus, despite the numerous obstacles faced through the experience of marginality, scholars in this field demonstrate that the collective consciousness and political agency generated within the spaces of education can also provide a public venue for social struggle and transformation. In Paulo Freire’s classic, *The Pedagogy of the Oppressed*, he notes that there is no such thing as a “neutral educational process,” and studies in geographies of education highlight its inherent political nature and corresponding spaces of contrasts and contradictions.

During the past few decades, the realm of education has faced innumerable challenges. These include the globalization and commodification of knowledge production; shifts in demographic composition and household structures spread unevenly across and within cities and regions; the retrenchment of the Fordist-Keynesian welfare state, affecting the redistribution of educational provision; and the neoliberalization of public policy reform, expounding and universally imposing its dictates of marketization, privatization, and accountability practices on the educational terrain. With the increased international migration and cross-border mobility in Western democracies, debates over multiculturalism, citizenship, and cultural identity rights; the unique needs of immigrant and refugee children; the provision of continuing education and second-language acquisition; the evaluation and negotiation of foreign credentials for adults; and the fluid circuitry of transnational experiences and outcomes have not escaped state-bound educational concerns. On a cautionary note, studies in geographies of education alert readers to the parallel hegemonic discourses involving the construction of new kinds of subjectivities that are constituted and entrenched spatially. Such discourses fuel a politics of fear, difference, and “dangerousness” and the targeting of “at-risk” students, which lead to further alienation of minority youth in particular. Landscapes of disparities have subsequently evolved at various scales, creating new spaces, subjectivities, and power relations within communities and schools.

Henry Giroux notes in *Theory and Resistance in Education* that changing historical conditions posit new problems, define different projects, and often demand fresh discourses. In recent years, empirical studies in geographies of education, particularly in critical geographic thought, have addressed the globalization of education reform and comparative educational policy; the planning, restructuring, and governance of educational institutions; the uneven distribution of educational provision, outcomes, and resources; and the politics of school closures and communities in decline. They have also included provocative debates over private versus public schooling, including religious, linguistic, and minority-focused schooling. Studies on neoliberal governmentality of education have examined the impacts of geosurveillance, new forms of regulation and accountability practices, volunteering and fund-raising expectations, and the proliferation of educational quangos (quasi-governmental organizations) within a wider culture of audit.

The complexities of educational systems have invariably led to broader theorizations and critical debates related to space, place, and society. For example, research related to citizenship, pedagogy, and technologies of government; the politics of race, identity, and integration; and spaces of exclusion, power, and negotiation are among the many converging themes emerging in recent scholarship on the geographies of education.

*Ranu Basu*

*See also* Children, Geography of; Citizenship; Democracy; Geography Education; Inequality and Geography; Knowledge, Geography of; Neoliberalism; Public Space; Segregation and Geography; Social Justice; Social Movements; Spatial Inequality

**Further Readings**


EGENHOFER, MAX (1958–)

Max J. Egenhofer is a geographer whose research has included work on spatial reasoning and knowledge representation; user interfaces for geographic information systems; the design of spatial database systems; and mobile spatial information appliances. He has been an influential figure in the development and analysis of geographical information systems (GIS) not only in academia but also in the private sector.

Egenhofer is a professor in spatial information science and engineering (since 1999), a cooperating professor in computer science (since 1991), and former director of the National Center for Geographic Information and Analysis at the University of Maine (1993–2006). He has held the Libra Professorship of the College of Engineering at the University of Maine and visiting professor appointments at the Westfälische Wilhelms-Universität Münster, Germany (2004), and the Università de L’Aquila, Italy (1993). He earned an MS in surveying engineering from the University of Stuttgart, Germany, in 1985 and a PhD in surveying engineering from the University of Maine in 1989.

Egenhofer has proved to be an exceptional research leader. As director of the National Center for Geospatial Intelligence Standards for 12 years, he made this unit the premier place worldwide for research in the technical domain of geographic information science, and he moved the University of Maine into the top rank among universities that receive research funding from the National Geospatial-Intelligence Agency. His service to the scientific community includes the organization of more than 20 national and international conferences. He is the founder of a highly successful GIScience conference series, has been program chair of most academic conference series in GIS, is a past editor of GeoInformatica, sits on six editorial boards, and has served on more than 120 program committees.

Egenhofer’s work has had a significant influence in industry. Several international standards have been formulated explicitly around his work, and a series of commercial GIS and spatial databases (by Oracle, IBM, Microsoft, Intergraph, Environmental Systems Research Institute [ESRI], Intelligent and Spatial Technologies) have implemented his formalizations. In ISO 19107, the model of topological relations he developed was the “Egenhofer Operators.” Fittingly, the University of Maine recognized him with the 2002 Presidential Research and Creating Achievement Award, and the UCGIS has honored him with the 2003 Researcher of the Year Award.

As an educator, Egenhofer has had an enormous impact on generations of doctoral and master’s students in SIE. His SIE graduates have successfully published in competitive outlets and have gone on to become leaders in academia around the world or head their own successful companies (e.g., IST, Planet One, Spatial Minds).

Egenhofer has published more than 100 articles in refereed outlets, among which are the most frequently cited papers in four journals and another three conference series, including the most cited paper in the GIS flagship publication,
the International Journal of Geographical Information Science. Egenhofer’s research record includes more than 50 research grants from the National Science Foundation, the National Institutes of Health, the National Geospatial-Intelligence Agency, the Federal Geographic Data Committee, the U.S. Air Force Research Laboratory, the Advanced Research Projects Agency, the Central Intelligence Agency, ESRI, and many others.

Maria Vasardani

See also Geographic Information Systems; GIScience

Further Readings


ELDERLY, GEOGRAPHY AND THE

In the absence of migration, the population of any area will become older when fertility and mortality levels decline. Western industrialized nations have already moved into the phase of pronounced population aging in the face of both fertility declines since the mid 1900s and an increase of almost 30 years in life expectancy during the 20th century. Over the next three decades, the aging of a large post–World War II birth cohort will cause an acceleration of this aging trend. Despite the overwhelming focus on aging in Western societies, the vast majority of the world’s older individuals reside in less developed countries outside North America, Europe, and Japan. Rapidly changing political, economic, and social conditions within these areas are likely to result in precipitous shifts toward lower mortality and fertility. With the modernization of these societies and a trend away from reliance on family-based care, the pending demographic tsunami is anticipated to create major problems in providing adequate care for elderly populations as they become frail.

Changing Location and Migration Patterns

The aging of economically developed societies in association with population redistribution (involving the migration of elders and labor force populations) has resulted in a geographical concentration of older populations. Regionally, patterns of aging-in-place lead to concentrations of elders in rural areas abandoned by out-migration of the young in search of employment. Increasing concentrations of retiree migrants are to be found in coastal and amenity-rich communities. The international migration of elderly populations is also becoming an increasingly significant phenomenon. On the scale of individual cities and communities, urban core areas have historically experienced high concentrations of elders as a result of abandonment during the suburbanization of the young. Many rural communities, especially small towns providing attractive retirement destinations, also have high concentrations of elders. Such patterns are complemented by a trend toward the spatial separation of older populations, both within naturally occurring retirement communities and in planned age-segregated retirement communities ranging from small groups of elders housed in specially designed housing complexes to communities of more than 75,000 elders, such as Sun City, Arizona.

Elders’ Experience of Space and Place

Burgeoning elderly populations have led to an increasing concern with understanding how aging people use space and experience place and the manner in which such experience evolves as a result of changing physical, social, and psychological capabilities associated with advancing years. Older people spend more time at home. Limited mobility leads to progressively more constricted activity spaces. This pattern may be associated with an increasing emphasis on vicarious environmental participation and a growing
emotional attachment to easily negotiated and familiar settings including the immediate residence. A desire to preserve independence and age-in-place may be manifested in a geographical centralization of activity within a limited area within the home. Elders may engage in the proactive creation of place by surrounding themselves with treasured artifacts, including photographs that sustain an ongoing sense of home and personal identity. This, in combination with the physical and psychological advantages of environmental familiarity, may result in a reluctance to relocate.

Geographical Aspects of Service Provision and Care

Creation of a continuum of care involving an ever-expanding array of services and environments represents one societal response in most Western countries to the growing numbers of elders. In the United States, the public sector of this continuum is coordinated through a geographic network, Area Agencies on Aging, created under the aegis of the Older Americans Act of 1965. A pervasive imperative to preserve elders’ independence and ability to age in place has given rise to home-delivered services (e.g., home care), enhanced community-based resources (e.g., nutrition programs and senior centers), and an ever-widening array of supportive residential options including sheltered housing, assisted living, and continuing care retirement communities. At the end of the continuum are nursing facilities providing skilled medical care to the most frail and hospice alternatives for those in the final phase of life. Major challenges are presented by questions of geographical accessibility (location of facilities within communities and relative to family networks) and social access (cultural expectations and ability to pay) to the continuum. As elders move along the residential continuum, the issue of relocation and associated influences on morbidity and mortality becomes increasingly significant.

Emergent Environments of Aging and Enduring Policy Dilemmas

The graying of populations has had major impacts on the geographical landscape over the past century through the transformation of community spaces and built environments as successive generations have reshaped the milieu in which people grow old in parallel with changing societal attitudes toward elders. It has also given rise to a set of enduring issues with significant geographical implications. What constitutes successful or optimal old age may change as priorities focus more on aging in place and maintaining independence. A growing trend toward geographical separation, in some cases the segregation of older populations, must be balanced with the benefits of social integration across the population age spectrum. There is a further need to identify the appropriate role of families versus societal institutions in maintaining the personal autonomy of older people. Reconciling age and need as criteria for the provision of services represents an ongoing challenge for the welfare state. Resolution of such issues is central to the creation of spaces and places that will ensure optimal aging.

Graham D. Rowles and John F. Watkins

See also Demographic Transition; Gated Community; Health and Health Care, Geography of; Housing and Housing Markets; Migration; Population Geography; Population Pyramid

Further Readings


Electoral geography is the study and analysis of the geographic patterns of elections, referenda, and legislative votes. The field originated in the early 20th century, particularly with the work of the French scholar André Siegfried, who produced some of the first maps showing the percentage of votes for parties by local region. Traditional electoral geography has focused on such spatial patterns of votes. Since the early 1970s, however, the field has also included other concerns such as the “neighborhood effect” and the spatial dynamics of political life.

Traditional Electoral Geography

Producing a map of election results is the most basic technique in electoral geography. These maps reveal the relative strength of support for candidates, parties, or referenda in different regions and are now often a prominent feature of election coverage in the popular media. The infamous red-blue map of the United States from the 2004 election, for example, showed Republican support stretching across a vast area of the country, while Democratic strength was concentrated in the northeast, the west coast, border regions, and scattered counties in the Midwest (Figure 1). On the basis of this map, Republican candidate George Bush appeared to enjoy broad support across the country.

Yet such simple maps can mislead in at least two ways. First, using only two colors disguises variations in strength among regions, with areas won by a landslide appearing the same as counties won by narrow margins. Better cartographic techniques—using gradations of color, for example—can overcome such problems.
The second, and more fundamental, problem is that such maps represent area rather than votes or population size. In this case, the map appears to be almost overwhelmingly red, although President Bush won by less than 2.5% of the popular vote (3.3 million votes nationally) and by just over 6% (34) of electoral college votes. Using a cartogram where the sizes of counties correspond to their populations produces a map with areas of red and blue that are more nearly equal but at the cost of distorting the shape of county, state, and national boundaries (Figure 2).

Despite these limitations, such descriptive maps or cartograms can provide a vivid and easily understood snapshot of political patterns for particular elections. Comparing a series of maps based on several elections, or a single map based on data from a series of contests, can reveal relatively stable political patterns, as well as areas of change. Particularly stable patterns can be used to identify and classify regional political cultures. Daniel Elazar, for example, divided the United States into three political cultural regions: Moral, Individual, and Traditional. These regions tend to consistently vote differently than each other over a series of elections, suggesting that each area has a different political culture.

Similarly, spatial patterns of voting are also used to trace the emergence and evolution of political parties. Parties tend to be more diverse and less stable in countries that use party-list systems (where voters cast a ballot for a party rather than for a particular candidate). When old parties decline and new ones appear, electoral geographers trace the way in which voters transfer membership and loyalty or, in some cases, discover

Figure 2  Cartogram by population of counties won by President George Bush (Republican candidate) and Senator John Kerry (Democratic candidate) in the 2004 U.S. presidential election
which voters display little party loyalty and are willing to switch between elections. Such transitions are typically tied to changes in regional economies, demographic and migration trends, and regionally based public investment. During the 1990s, there was special interest in the emergence of political parties in the democratizing states of the former Soviet bloc. In these cases, support for communist parties (or their direct heirs) or for extreme nationalist parties was often strongest in regions that had experienced more difficult transitions to a market economy.

Electoral Geography From a World-Systems Perspective

The basic analyses discussed above generally assume (explicitly or implicitly) that spatial voting patterns are relatively stable (the same or similar voters will vote in the same or similar ways over a series of elections), that voters consistently vote for political parties that best represent their interests, and that the resulting government policies will reflect the interest of the majority of voters. All these assumptions have been challenged by electoral geographies based in a world-systems perspective. Scholars using this framework—most prominently Peter Taylor—have argued that these assumptions only hold true (at best) in “core” states (wealthy, industrialized countries, primarily in North America and Western Europe). Electoral patterns outside the core tend to be highly unstable, as do elected governments in these regions.

There are two basic explanations for this difference. First, the modernization process that accompanied the development of representative democracies in Western Europe coincided with the emergence of two stable sets of party cleavages. Nationalization produced a split between subjugated and dominant cultures and between religious (church) and secular (state) authority. Simultaneously, industrialization produced a cleavage between agriculture and industry, and between capital and labor. These fissures remained stable in Western Europe through the late 19th and early 20th centuries, producing relatively stable spatial patterns and relatively stable party organization. Regional patterns of voting can be explained by alliances between different groups aligned along these two sets of differences. Other regions of the world, particularly those previously colonized by Europe, experienced a different modernization process that did not produce such stable political divisions. Second, world-systems theory suggests that parties remain in power only to the extent that they can form governments that effectively redistribute resources (i.e., social welfare services) to the electorate. Noncore states, however, are not wealthy enough to consistently provide enough resources to effectively cultivate the loyalty of constituencies. Consequently, elections outside the core are characterized both by low party loyalty and by the consistent electoral failure of incumbent parties.

Redistricting and Roll-Call Analyses

The practice of redistricting is a special case of electoral geography. In political systems that use election by district—where candidates run for office in a specific geographic constituency—the location and behavioral pattern of voters are often of concern for the organization responsible for drawing new districts. Political parties attempt to use redistricting to distribute their likely voters so as to create as many districts as possible in which the party has a majority or plurality. This process thus involves a kind of “applied” electoral geography, where the spatial distribution of voters and the likelihood that they will vote in particular ways (for or against particular parties or candidates) are used to influence the outcomes of elections. Scholarly analyses of various districting plans employ similar techniques, albeit without explicit political agendas.

Another subset of traditional electoral geography examines votes by representatives in legislatures and parliaments (so-called roll-call votes). Such analyses may focus either on regional variations in voting patterns or (more typically) on the relationship between the composition of constituencies and the votes of the corresponding representative. Figure 3 shows the relationship between the vote on an amendment to the Voting Rights Act, the party of congressional representatives, and the racial composition of congressional districts in Pennsylvania in 2006. In this case, one might hypothesize that representatives from districts with a significant number of African Americans—the traditional
beneficiaries of the Voting Rights Act—would have opposed this particular amendment (which would have weakened certain provisions of the legislation). Yet the map illustrates that at a minimum there are also important partisan effects at work: All Democratic and some Republican representatives (none of whom had a significant number of African Americans in their districts) also opposed the amendment.

The relationship between roll-call votes and independent variables such as party membership and constituency characteristics is often complex and typically needs to be modeled with multivariate regressions. Such studies are especially useful in examining the role of political representation in mediating the relationships between political preferences and political action. Unlike election results, roll-call votes reveal how particular individuals (representatives) voted, rather than simply presenting the aggregate results of many individual, anonymous voters.

**Figure 3** The geography of a 2006 congressional roll-call vote to amend the Voting Rights Act, by party, of the representatives and district percentage of African Americans in Pennsylvania

*Source: Author.*

**Scale and Electoral Geography**

Starting in about the late 1960s and early 1970s, electoral geographers began to move beyond traditional descriptive maps to use election results to investigate the spatial dynamics of political
behavior. These efforts were distinguished from world-systems theory in the sense that they examined the effects of space and spatial organization on elections and politics. Two directions in this endeavor are especially notable: studies of the “neighborhood effect” and political party organization and dynamics, and “place-based” studies of politics.

Studies of the neighborhood effect attempt to explain local spatial variations in voting patterns that cannot be accounted for by individual economic, social, and demographic characteristics typical at the national level. They hypothesize that voters who are similar but who live in different locations will vote differently if their local political contexts are different. Differences in such contexts can arise from at least two sources: acquaintance circles (groups of people who typically socialize with each other) and the organizing efforts of political institutions, especially political parties. In the first instance, people’s political preferences will tend to converge as they spend more time talking with each other, particularly in face-to-face situations. If a majority of a voter’s neighbors and acquaintances are conservative, for example, she or he may tend to become more conservative over time even if she or he otherwise has the characteristics of a liberal. In the second case, political organizing on the local level affects both voters’ perception of their own political interests and their likelihood to vote.

Finally, certain political geographers, most notably John Agnew, have advocated a place-based approach to the study of politics that calls for a substantial reworking of traditional electoral geography. Like “neighborhood” studies, such place-based or contextual approaches also examine spatial variations but do not automatically set “neighborhood” patterns against national ones. This method sees the aggregate nature of electoral data as an advantage rather than a limitation because it focuses on the social nature of politics and political identities rather than on individual preferences. In particular, political life and political interactions are seen as being inherently place based, with a web of preferences, social divisions and constraints, associations (including political parties), and institutions being organized and experienced in particular spatial contexts. The choice of scale for electoral analysis is fundamental, with national- or even regional-level electoral studies often disguising critical political relationships. From this perspective, microlevel electoral data (at the county, commune, or even precinct level rather than at the national level, for example) offer a better opportunity for analysis not because the data are more detailed (as traditional electoral geography might argue) but because the spatial units may correspond better to the scale of lived political experience. The approach does not assume a priori that there is a single best scale of analysis for electoral geography but rather that the appropriate scale depends on the specific situation and relationships in question.

Benjamin Forest

See also Agnew, John; Citizenship; Democracy; Johnston, R. J.; Political Geography; Redistricting; Scale, Social Production of; Taylor, Peter; World-Systems Theory

Further Readings


Electronic atlases are intentional combinations of maps or data sets structured in such a way that given objectives are reached. In the recent history of cartography, atlases deserve a prominent status among cartographic products due to the variety of their thematic content, the rather friendly appearance, and their utilization by a wide spectrum of users. Advances in information technology and more specifically in geographic information systems and digital mapping, have altered the traditional way of using maps and thus have a considerable impact on every aspect of electronic atlas design, content, and functionality. Considerable effort has been made in the production of electronic atlases, and successful systems are available with fascinating design and functional characteristics.

One way of understanding our neighbors better is to compare them with ourselves. The visualization of geostatistical data in electronic atlases with the support of basic analytical tools is the most effective way to represent quantitative and qualitative aspects of spatially distributed phenomena as well as their thematic and geographical relations.

Menno-Jan Kraak and Ferjan Ormeling consider atlases to be a higher form of cartography compared with maps and charts because their production involves both an extra planning and an extra structural dimension. Initially, electronic atlases emulated the appearance of paper atlases and were sequential “slide shows.” The more advanced electronic atlases have high interactive mapping capabilities that allow users to manipulate atlas data, to combine them with their own, and to examine spatial and thematic data dynamically, thus developing different representations. The distinction of electronic atlases into view only, interactive, and analytical is only an indication of the course followed until we reached the point where an electronic atlas is expected to virtually cover all realistic needs of the user community; from changing color schemes and the number of classes to combination, query, analysis, and dynamic presentation of geographic and statistical data sets. On the other hand, the increasing demand for sophisticated data presentation and manipulation is further stimulated by developments in scientific visualization, graphic user interface design, cognitive—usability, multimedia, and exploratory data analysis.

Another consequence of the technological change is the ability to access maps and related information via the Internet, which has changed people’s daily lives. The Internet can serve as a source of spatial data, thematic data, finished maps, and software for exploring and analyzing data. Despite the increasing trend of accessing, processing, and displaying information via the Internet, a well-designed electronic atlas should not be restricted in this mode of operation. The system should enable the user to be capable of using the electronic atlas as a stand-alone—where the atlas will be stored on a removable medium such as a DVD—supported by an Internet alternative, which will provide the updating of the geographical and statistical data.

Lysandros Tsoulos

See also Cartography; Internet GIS; Multimedia Mapping

Further Readings

ELECTRONICS INDUSTRY, GEOGRAPHY OF

The electronics industry manufactures the suite of information and communications technologies that harness the processing capacity of semiconductors, or microchips. In practical terms, the core set of electronics industries includes chip design and fabrication, computers, consumer electronics, and communications equipment (e.g., cell phones). The boundaries of the sector could reasonably be drawn to include a range of associated industries such as advanced computer-numeric controlled machinery producers or robotics (though not software) industries that rely substantially on electronic engineering. Advances in electronics have had a pervasive impact on virtually all facets of economic activity over the past 50 years. Given this almost singular importance, the industry has proven especially crucial in driving regional development in the places where it locates. The following paragraphs discuss the industry’s evolving spatial dynamics by emphasizing the roles of states, lead firms, and informal routes to skill formation.

Several key patterns define the geographical evolution of the electronics sector. The United States and in particular key regions such as Massachusetts’ Route 128 and California’s Silicon Valley are acknowledged as the crucibles of the electronics revolution, their emergence coinciding with the first commercial applications of the transistor in the 1950s. Silicon Valley persists as the world’s leading electronics industrial district. It is home to key firms such as Intel, Hewlett-Packard, and Apple and derives additional advantage from its proximity to Stanford University, the sine qua non model for university technology transfer. The positive externalities of Silicon Valley act as a magnet for American and international talent, making it virtually imperative for the world’s leading electronics firms to have a presence there too.

Apart from the United States, Japan has been perhaps the biggest foreign beneficiary of the early advances in American electronics technology. Indeed, the transfer of American technology and basic science to Japan, most famously exemplified in Sony’s purchase of the Bell Labs/Western Electric transistor technology in 1954, provided the foundations for the Japanese consumer electronics firms, whose brands continue to dominate the industry. More generally, the emergence of Japan signaled the ascendance of East Asia as the center of gravity for electronics manufacturing.

The diffusion of the electronics sector within East Asia reflects shifts in the industry’s increasingly elaborate division of labor. Most basically, the electronics industry can be categorized into, on the one hand, high-value-added and specialized functions, such as chip design/production, that require significant economies of scale and state support—capital costs for new chip fabrication plants run upward of $3 billion to $4 billion. On the other hand, more routine functions such as component assembly for consumer electronics are attracted to low-labor-cost locations. An overwhelmingly female labor force persists as the constant characteristic of factories in these latter contexts.

The precise spatial articulation of this division of labor results from the twin motivations of firms and states. Electronics firms generally conduct higher-order functions such as R&D and pilot manufacturing in their home markets. This means that most R&D and patenting in electronics are concentrated in a relatively narrow set of advanced industrial nations with significant basic scientific capabilities (the United States, Japan, Germany, Finland, South Korea, etc.). In contrast, the routine production of electronics that occurs at later stages in the product cycle almost inevitably takes place in lower-cost locations. States that host multinational electronics firms therefore have a powerful incentive to enhance the skills of their workforce so that they can capture the investment of increasingly more sophisticated functions that are likely to spawn a domestic electronics sector.

To illustrate this crucial firm-state relationship in East Asia, Japanese electronics firms, for example, have exhibited a sequential pattern of investments through a succession of host nations, including Taiwan (1960s); South Korea and Singapore (1970s); Malaysia, Indonesia, the Philippines, and Thailand (1980s); and China and Vietnam (1990s). By using Japanese investments as a catalyst for learning, South Korea, Taiwan, and more recently China have been successful in cultivating domestic R&D-driven electronics industries capable of higher-order manufacturing/design functions. In contrast, Malaysia, Indonesia,
the Philippines, and Vietnam have served principally as magnets for investments in assembly-type functions. The lesson to be drawn from this example is that the economic development potential from electronics investments clearly hinges on the ability of states to both bargain for higher levels of technology transfer and encourage domestic skill formation for the management of more advanced functions. While formal state support and large firm investments are essential conditions for fostering an electronics sector, other less formalized factors are also important.

The success of the Japanese electronics industry in the postwar era owes much to the more informal skill base prevalent in Japan at the time. A case in point is the genesis of Akihabara, Tokyo’s retail Mecca for electronics and pop culture content such as anime and video games. Following World War II, a network of amateur tinkerers descended on Akihabara’s weekend bazaar, a former black market, for surplus military radio components. With guidance from electronics digests, this population learned to engineer their own radios and similar proto-electronics. As the Japanese electronics sector took off in the 1950s and 1960s, it leveraged the tacit skill base of this tinkerer cohort. Steve Wozniak, Apple’s cofounder, tells a similar tale about his youth in Silicon Valley. Perhaps these historical anecdotes are a prologue, as poorer nations such as China and Bangladesh, which currently serve as the destinations for the world’s cast-off electronic waste, may be able to translate the skills of today’s component recyclers into a more formal electronic industry in the future.

Tim Reiffenstein

See also Flexible Production; High Technology; Industrial Districts; Innovation, Geography of; Product Cycle; Research and Development, Geographies of; Technological Change, Geography of

Further Readings


EL NIÑO

The El Niño-Southern Oscillation (ENSO) is a joint oceanic (El Niño) and atmospheric (Southern Oscillation) phenomenon spanning the equatorial Pacific and is one of the more important causes of global climate variability. Its first-order impacts are on the distribution of sea surface temperatures (SSTs) and precipitation over the equatorial Pacific and environs. These changes in the locations of the vast quantities of energy stored in the ocean waters and their release to the atmosphere by means of latent heat during cloud formation are spread around the globe by a series of atmospheric linkages known as teleconnections. Differences in the speed of response of the atmospheric (rapid) and oceanic (slower) components of this coupled system ensure that it is intrinsically unstable, oscillating in a pseudoperiodic fashion with a frequency of approximately 3 to 7 yrs. (years) between three principal phases: “warm,” closely linked to the popularly named “El Niño”; “cold,” also variously called “La Niña,” “El Viejo,” or “Anti-El Niño”; and “neutral.” The nomenclature refers to the dominant sea surface temperatures in the central and eastern equatorial Pacific rather than the resultant
local and regional climate conditions, which can be warmer or cooler, and wetter or drier, depending on the season and the geographic location.

**History of Discovery**

The term *El Niño*, or “Christ child” in Spanish, has been used for centuries by fishermen of the Northern Peru and Southern Ecuador littoral to describe a quasi-annual warming of coastal ocean waters shortly before Christmas, which persists for several months. Cold-water pelagic fish, such as anchovies and sardines, dominated and provided the basis of the regional maritime ecology, including guano birds, whose droppings constituted a large portion of the global fertilizer market, and the Peruvian economy before the advent of chemical fertilizers. The cold waters are the northernmost expression of the Chile, Peru, or Humboldt Current, flowing northward along the coast of Western South America from the sub-Antarctic Ocean and westward along the equator in a semipermanent feature called the “cold tongue.” Fishermen noted this seasonal change in the waters through the fluctuating size and composition of catches. Tremendous variability in the severity, onset, and duration of this phenomenon was also noted. On some occasions (“cold phase”) the losses were negligible, while other years (“warm phase”) brought a total collapse of the fishery and extensive deaths of guano birds and marine mammals. Colonial records from Peru indicate that the impacts of these changes were often exacerbated by contemporary devastating floods in the coastal Chira desert region and simultaneous droughts in the Altiplano.

Following a sequence of severe droughts, failed harvests, and millions of deaths resulting from the failure of the monsoon rains in India during the late 19th century, a leading British atmospheric scientist, Sir Gilbert Walker, was dispatched to determine their cause. Examining meteorological records from the subcontinent and surrounds, Walker noted a strong regional association in annual variations of atmospheric pressure over the area encompassing southeast Asia, particularly the Indonesian archipelago, and bordering the western equatorial Pacific and Australia, then another British dominion that was subject to severe and devastating climatic variability. Encouraged by this hemispheric-scale association in atmospheric pressure variations, Walker extended his analysis to the global scale and noted a strong negative correlation between the atmospheric pressure over Indonesia and Australia (Djakarta and Darwin) and the Central Pacific (Tahiti). Interannually, pressures above these two nodes appeared to rise and fall in lockstep with each other but in opposing directions. Walker coined the term *Southern Oscillation* to describe this observation. Further analysis of fluctuations in pressures at other locations around the world offered hope that the phenomenon might provide an explanation for variations in what Walker termed *global weather*.

At the time, no connection was drawn between the oceanic El Niño phenomenon of the Eastern Pacific and the atmospheric Southern Oscillation in the west. Ironically, it is now clear that shortly after the publication of Walker’s work in the late 1920s and early 1930s, interest in the linkages in his “global weather” entered three of four decades of relative quiescence, till after his death in 1958. His obituary alludes to a prevailing contemporary feeling that Walker’s belief in the global significance of the Southern Oscillation was misplaced. However, within a decade of his death, a resurgence of interest in global fluctuations of climate and fish stocks and the reemergence of stronger “warm-phase” events (1965–1966, 1972–1973) led to a reevaluation of Walker’s contributions and the positing of causal linkages between oceanic and atmospheric phenomena as a means by which impacts are spread globally. In the past 50 yrs., scientific and public awareness of the significance of Walker’s discoveries and the impacts of ENSO have expanded greatly, aided by the devastating El Niño (warm-phase) events of 1982–1983 and 1997–1998 and a reevaluation of its role in global environmental, social, and political history.

**Physics of ENSO**

Warm air rising above the equatorial oceans causes surface winds to move in from both south and north of the equator to replace it. At the top of the troposphere (8–10 km [kilometers]), the air moves poleward, cooling, to descend at approximately 30° north and south. This continuous exchange of air north and south of the equator at
various altitudes is termed Hadley Cell Circulation. Air masses are deflected to the right (Northern Hemisphere) and left (Southern Hemisphere) of their intended path by the Coriolis force, resulting from Earth’s rotation. The resultant surface winds, the northeast and southeast trades, are some of the most noticeable and permanent features of the global wind fields. The two winds converge in a broad discontinuous band of warm rising equatorial air known as the Inter-Tropical Convergence Zone (ITCZ). The ITCZ migrates seasonally, driven by the position of the overhead sun, differences in the thermal properties of oceans and continents, the global distribution of continental masses, and the underlying SSTs. The net result in the Pacific is to convey warmer surface waters away from western continental South America, along the equator toward Australia and southeast Asia, leaving a “warm pool” of warm water in the western ocean and a cold tongue in the east.

A similar geographic disparity is observed beneath the ocean waters. In the west, accumulated warm surface waters physically raise the ocean height and also drive the boundary between the warmer surface and cool subsurface waters, the 20°C thermocline, to 150 m (meters) beneath the surface. Along the Peruvian coast, cold subsurface waters appear at the surface, and the thermocline is located at a much shallower depth (40–50 m). The atmosphere above these two areas responds quickly to the differing SSTs, intensifying the low atmospheric pressures and lifting of air above Indonesia (SST 28°C), while the cool waters along the cold tongue (20–22°C) produce cooler, more dense, subsiding air. The air rising over Indonesia produces excessive rains and a transfer of latent heat to the atmosphere, while clear skies prevail in the eastern ocean. The rising air spreads east and west at the top of the troposphere, contributing, in part, to the descending air over the cold tongue. This cell of east-west circulating air and ones like it above the equatorial regions are termed Walker Cell Circulation, commemorating the seminal contributions of the climatologist.

At some point, the strength of the trade winds and the north-south Hadley Cell weaken, allowing the warmer water to migrate westward. More important, a wave motion is set up in the thermocline, which travels faster than any surface waters, spreading eastward across the Pacific with concomitant changes in SSTs. Such changes are generally detectable about the international date line in June or July, with their impacts reaching coastal South America in late November or December. As the wave moves eastward, it cannot dissipate north or south because the Coriolis force focuses the phenomenon along this equatorial belt. The precipitation associated with the warm pool moves eastward to the Central Pacific, producing a drought over Australia-Indonesia. This shift in associated release of energy in the form of latent heat strengthens the north-south Hadley Cell circulation over oceanic regions, enhances the trade winds, and begins to reverse the process. Generally, about 12 to 18 months elapse between the first appearance of warmer waters in the Central Pacific and the conclusion of a “warm-phase” event.

The cessation of a warm phase often portends intensification of both the warm pool and the cold tongue (“cold phase” of ENSO). The east-west Walker Cell now dominates, and the Hadley Cell begins to weaken, bringing the process back to its original condition with the noted periodicity. In this way, the conditions of both cold and warm phases ultimately contain the seeds of their own demise and tend toward the development of the opposing phase because of the different response times of the coupled atmospheric and oceanic systems. As such, the arrangement is inherently nonlinear and displays many of the properties of a chaotic system, with three nodes of attraction or the three phases of ENSO. No pair of similar phases is ever the same in its development, intensity, or duration, making the forecasting of the nature and extent of impacts difficult and sometimes leading to differences of categorization phases, which is actually all part of a continuum.

Global Impacts of ENSO

The direct impacts of ENSO are felt along the equatorial belt of the Pacific and marginal continents. In these regions, extra cloud cover is generally associated with cooler air temperatures and vice versa. Warm phases bring drought to the Western Pacific, as exemplified by the extensive forest fires of 1997–1998 in Indonesia. Excess rains fall over the Central Pacific islands as the
warm pool migrates and the weakening of the descending limb of the Walker Cell circulation over coastal Peru, combined with the diminished upwelling of cold water and a general warming of the near-shore waters, permits the progression of the ITCZ south of it general southernmost position of 3° N, bringing torrential rains to the coastal desert region. Figure 1 illustrates the diverse climatological impacts of ENSO around the world. Their exact timing depends on the interaction of the local annual cycle and that of ENSO. The strength and effectiveness of the teleconnections vary from event to event, with local conditions, notably the SSTs of the Indian, North Atlantic, and South Atlantic Oceans, either amplifying or dampening the transmission of the ENSO signal.

**Long-Term Changes**

Within the period of historic instrumental records, there is evidence for lower-frequency (decadal to multidecadal) changes in these teleconnections. A shift in global climate in the 1970s changed circulation patterns in the North Pacific; changes in the Indian Ocean are believed to change associations with the South Asian monsoon and the climate of East and southeast Africa; and changes in the temperatures in the tropical North and South Atlantic control signals to northeastern Brazil and West Africa. Spanish colonial records from Peru and Chile indicate that ENSO has been operating along this coast for more than four and a half centuries, while archaeological, ice core, and lake
sediment evidence indicate the persistence of extreme ENSO phases over at least 1,000 yrs.

Peter Waylen

See also Atmospheric Energy Transfer; Atmospheric Moisture; Atmospheric Pressure; Atmospheric Variations in Energy; Coriolis Force; Hadley Cell; La Niña; Oceanic Circulation; Teleconnections

Further Readings


Emerging Markets

The term emerging markets typically refers to a distinctive financial asset class among international investors and financial firms and designates countries perceived as transitioning from developing to developed economies (often countries that were formerly part of the Third World or the communist bloc). The term emerging markets was coined in 1981 by Antoine van Agtmael, the then director of the Capital Markets Department of the International Finance Corporation (IFC), the private sector development arm of the World Bank Group, which created emerging markets as a financial asset class when it developed the first emerging market investment fund (the Korea Fund, in 1984) and provided the seed money for that fund. Between 1984 and 1994, the IFC underwrote and invested in more than 30 emerging market portfolio investment funds before spinning them off to private investment firms. There are now more than 600 equity emerging market investment funds, with most of the capital originating from international and domestic U.S. funds, offshore accounts, hedge funds, and insurance companies.

The countries designated emerging markets are not fixed as the criteria for defining them vary among investors and financial firms. However, the following IFC/World Bank criteria are typically used in identifying emerging market countries. They have experienced 3 yrs. (years) or more of rapid economic growth (based on gross domestic product [GDP] figures) yet remain low- or middle-income economies according to the World Bank. They have implemented liberal economic policies and practices such as adopting international financial standards and broad-based discriminatory controls for nondomiciled investors. They do not exhibit financial depth (i.e., the ratio of the country’s market capitalization to its GDP is “low”) and hold high levels of foreign debt and are subject to currency risks. They have a functioning stock market but often inadequate regulatory oversight, lack of transparency and liquidity, and broker, settlement, and custodian risks. Finally, emerging markets are typically considered risky investments because they are countries often considered economically, institutionally, and politically unstable.

Emerging markets represent more than 12% of world market capitalization. Overall private capital flows to emerging markets, as a percentage of their GDP, are now more than four times net official aid flows to those same nations. The issuance of international securities by emerging market sovereigns and corporations has increased from a level of $325 million in 1995 to roughly $700 million by 2003. By 2006, emerging market securities had increased in value by 243% compared with the United States, which had increased only 34%. Net equity flows (foreign direct investment and portfolio flows) to emerging markets have grown
to roughly $200 billion per yr. By 2006, emerging markets (led by China, Taiwan, and Korea) held more than 75% of all foreign exchange reserves, while the largest emerging markets (the so-called BRIC countries of Brazil, Russia, India, and China) now hold foreign reserves totaling more than double their foreign debts. The holding of foreign currencies is considered indicative of high levels of international trade and market integration. Economic integration has also produced a heightened tendency toward financial volatility and crises as the frequency of financial crises in emerging market economies since 1973 has become twice that of the Bretton Woods period.

There are currently more companies listed on emerging market exchanges than there are companies listed on developed country exchanges. The most common financial instruments used in emerging market funds include American Depository Receipts (ADRs) and closed-end country funds (number of shares in the fund are fixed). ADRs are popular because they circumvent many of the geographic problems involved in international financial transactions (such as avoiding currency conversions, use of domestic brokers, and clearing and settlement institutions) affiliated with many emerging market exchanges, while closed-end funds allow emerging market states to regulate capital flows. However, the number of open-end country index funds (number of shares changes) is increasing.

It has been suggested that labeling former communist bloc and semi-industrialized Third World countries as emerging markets is symptomatic of the diminishing credibility of a passé Cold War geopolitical division of the world. Emerging markets have become more than simply a standard label for the semiperiphery in the unfolding world system of the new millennium. It has been argued that emerging markets have been produced by a geofinancial consortium of global news networks, international stock market analysts, mutual fund managers, institutional information agents, economists, debt security and bond-rating agencies, United Nations observers, and foreign affairs intelligentsia that actively rescript global economic space to imply an increasingly integrated (or globalized) world financial system. The discourse surrounding emerging markets speaks of them as “emerging” from underdevelopment and economic backwardness while taking the proper steps up the ladder of development, and thus, emerging markets serve as models of development for the rest of the developing world to emulate. The media and financial sector literature is reminiscent of colonial discourse using metaphors of frontiers, pioneers, and exploration in places perceived of as wild and risky. Of course, the rhetoric also conceives of the world as a field for global capital flows and profit taking where emerging markets are treated instrumentally as yet another space for investing surplus capital, selling products, and extracting resources.

Jayson J. Funke

See also Developing World; Finance, Geography of; Globalization; Money, Geographies of; New International Division of Labor; Offshore Finance; Political Economy; Underdevelopment; World Bank; World-Systems Theory

Further Readings


Emotions have always had an important part to play in the theory and practice of geography. This observation holds despite what might be seen as a recent “emotional turn,” or persistent academic desires to keep such affects under scholarly wraps.

Among other reasons for their apparent absence from mainstream geography is the frustrating intangibility of emotions; they are impossible to
locate, map, or measure in ways characteristic of the geographer’s more typical field or object of study. Emotions are difficult to place but are nonetheless always everywhere. At times, they seem to come from “within” to affect the world around us, as when feeling “blue” all but blinds us to the beauty of a place that at other times (in other moods) can move us to experience joy. They can thus also seem “external” affairs; we may “find” happiness on the beach, find that graveyards “cause” despair, or find that we are bored by classrooms and rainy days that “make” us long to be elsewhere. Emotions are, then, relational, caught up with (in and around) people and places in profoundly important ways, and geographers are well placed to investigate such phenomena that shape and are shaped by the different—sociocultural, economic, political, and so on—worlds we live in.

One of the projects undertaken by recent works in “emotional geographies” has been to uncover emotion’s hidden histories in the discipline and to reveal its suppression as well as expression in our various ways of “writing the world.” Emotions have thus been shown to matter for geographies of all kinds, even those supposedly abstract or, say, economic. It is widely acknowledged that people’s feelings about and “sense” of place will have an impact on the most significant purchase many make during their lives and can help transform those mortgaged houses into comfortable, comforting homes. Considering larger spatial scales, global financial markets are notoriously sensitive to consumer confidence or fears, and a population’s aspirations and anxieties have long been seen as significant for political geographies too; the increasing presence and importance of “hope” in political discourse is becoming difficult to ignore, especially in the United States. Beyond such subdisciplinary concerns, emotions are clearly central for contemporary human geography—hardly surprising once we recognize how vital to human experience our emotions really are.

Recent geographical attempts to research and theorize emotions tend to coalesce around four main approaches or schools of thought. Feminist geographers such as Gillian Rose have been influential in questioning dualisms and insisting that all experience and knowledge is situated and relational, embodied and full of feeling yet never separate from or opposed to reason. Geographic phenomenologists such as Yi Fu Tuan and Edward Casey provide insights into perceptions of and emotional responses to particular places. Psychoanalytic/psychotherapeutic geographers, such as Liz Bondi and David Sibley, highlight the importance of the formation and maintenance of psycho-social boundaries for understanding the emotional intersections between selves and spaces. Finally, nonrepresentational theorists such as Nigel Thrift attempt to negotiate the challenges associated with articulating emotional relations and affects by focusing on performance, on what people do rather than what they say they do. While such approaches differ in many ways, their shared interests are substantial and help advance geographical understandings of emotions that motivate how and with whom we spend our time, how and where we live and work, and indeed what we, as geographers, choose to study.

Joyce Davidson

See also Behavioral Geography; Body, Geography of; Everyday Life, Geography and; Fear, Geographies of; Feminist Geographies; Hate, Geographies of; Humanistic Geography; Identity, Geography and; Nonrepresentational Theory; Phenomenology; Psychoanalysis, Geography and; Sense of Place

Further Readings


Empiricism is a long-standing and highly influential philosophy, particularly within geography, which begins with the assertion that the only valid form of knowledge is experience and sense perception. That which is not observable is held to be nonexistent or meaningless.

Although its origins may be traced to scholars such as Francis Bacon, empiricism was largely a child of the British Enlightenment, in which intellectuals such as John Locke, David Hume, and George Berkeley emphasized the primacy of logic and empirical reality in the face of medieval metaphysics—that is, evidence rather than faith; in this sense, empiricism was a progressive ideology in its day. British empiricists also deployed the philosophy in their debates with continental rationalists such as Descartes, Spinoza, and Leibniz, who held that knowledge came from innate ideas alone. Without sensations and experience, empiricists maintained, the mind would have nothing to work with.

At the core of empiricism is the notion that statements about the world, that is, claims regarding sensations and data, are true or false without appeal to theory. The word *data* derives from the Latin *dar*, "to give," indicating that data are given, not made. In this view, a fact is simply a statement about what is, an assertion of verified information that exists beyond reasonable doubt. (Ironically, the Latin *factus* refers to that which is made, not simply given.) There is thus an iron wall that separates facts and opinions, which empiricists dismiss as subjective. Facts—and the truths they lead to—exist independent of their interpretation and are thus true whether or not anyone knows or believes them, indicating that truth has an objective existence independent of the knower.

Empiricism shares some aspects with logical positivism, which also holds to a sharp dichotomy between fact and theory, the objective and the subjective, and a disregard for the role of the observing subject. However, unlike empiricism, positivism maintained a healthy regard for theory, particularly deduction, the view that if the premises of an argument are true, then so must be its conclusions. Deduction is thus the move from the abstract to the concrete; in this view, an observation is true because it consists of a special case of a broader principle. Empiricism, however, generally subscribes to inductive, not deductive, logic, that is, the attempt to move from the observed to the unobserved, from the concrete to the abstract, from the specific to the general, from pattern to prediction. The explanatory basis of empiricism is the attempt to form generalizations without explanations. Unlike rationalism or logical positivism, empiricism holds that all conclusions are only probable or tentative and cannot be known with certainty, introducing a degree of skepticism and humility. Moreover, empiricism offered a healthy distrust of theoretical overdetermination and armchair speculation.

Empiricism has been a long-standing tradition in geography; for most of the discipline’s existence, it consisted of little more than the encyclopedic collection of data. The philosophy reached its zenith during the heyday of chorology in the early to mid 20th century and particularly in the work of Richard Hartshorne. In this line of thought, geography was the integrative but static description of regions, and explanation was the province of other disciplines. Empiricist geography tended to celebrate the unique and idiosyncratic at the expense of the abstract and theoretical, and the discipline acquired a popular reputation for being an analytical dead end.

Not surprisingly, abstraction and causation have always been the Achilles' heel of empiricism. The great danger of induction is in making sweeping and unwarranted generalizations or in confusing coincidence with causation. As the social sciences became increasingly self-conscious epistemologically, it became abundantly clear that the fact-opinion dichotomy was a false and misleading one: All data are theory laden, in the sense that it is theory, even if it is vague and unspecified, that tells one what data are relevant and important and what data are not. There are no “pure” facts as observations are always couched in terms of a theory: Observation requires interpretation, which in turn requires theory. Theory is necessary to move from description to explanation, to establish causality, test hypotheses, justify arguments, and make claims to truth. Theories are thus indispensable to knowing how the world works; understanding theory is not a choice because theory is inescapable. Thus, as geography and other disciplines moved into a welter of different theoretically
sophisticated schools in the late 20th century, including Marxism, phenomenology, and postmodern/poststructuralist frameworks, empiricism lost its credibility and today is virtually extinct.

Finally, it is important not to confuse empiricism with empirical work; many theoretical perspectives collect and analyze data in different ways. Almost all schools of thought place some value on empirical work to test or complement theoretical propositions, but the simple-minded upholding of data for their own sake, and the associated distrust of theory, is the unique province of empiricism.

Barney Warf

See also Chorology; Epistemology; Hartshorne, Richard; Human Geography, History of; Idiographic; Logical Positivism

Further Readings


Energy and Human Ecology

Ecology is a holistic approach to studying how biological, hydrological, geochemical, and even social processes are shaped by interrelated relations between living beings and ecosystems. The term human ecology is probably most associated with the 1920s Chicago School of Urban Sociologists, who borrowed concepts from plant ecology (e.g., metabolism, competition, succession) to explain the growth and spatial patterns of American cities. In 1923, however, another University of Chicago scholar, Harlan Barrows, attempted to define the entire discipline of geography as “human ecology.” In contradistinction to the urban ecological school, Barrows did not want to simply use ecological metaphors to explain social patterns but sought to remake geography as a science of ecological relationships between human societies and the natural environment. For Barrows, geographers would examine all types of human livelihood patterns—from agriculture to industrial manufacturing—and trace the relationships with landscapes of natural resource production and evaluate the ecological impacts.

For reasons still debated, Barrows’s “human ecological” approach did not remake geography into a science of nature-society relationships. Barrows’s students—most notably Gilbert White—went on to undertake important policy-relevant research on the management of resources and natural hazards such as groundwater and floods. Yet none of this work explicitly invoked an ecological approach focused on mutually constitutive relations, flows, and networks between human and natural systems. Only 40 years or so after Barrows’s initial proclamation would geographers discover a concept that researchers could effectively operationalize into human ecological research—energy.

Energy as an Ecological Concept

In the 1960s, Howard Odum helped develop the “systems ecology” approach, which situated energy flows at the center of analysis. A self-proclaimed energetic determinist, Odum suggested that all natural and social systems could be
understood in terms of the energy available to accomplish work. Indeed, the very basis of life on Earth depends on plants making use of solar energy through the process of photosynthesis and respiration, which, in turn, provide energy for a variety of herbivores and omnivores, who, in turn, provide energy for carnivores. Moreover, since humans harnessed the concentrated sources of solar energy latent in fossil fuels (coal, oil, and natural gas), Odum suggested that energy is also at the basis of seemingly unnatural processes such as urbanization and industrialization. Through characteristic flow diagrams and circuit models, Odum and other systems ecologists mapped the ecological flows of energy and matter in a variety of systems from wetlands to households.

This systems ecological approach considers energy from the basis of the laws of thermodynamics. Although there are several debates on several laws, most agree that the first two are the most important. The first law states that from the perspective of an open system, energy can never be created or destroyed but only change forms. Ecologists can trace the different forms of energy from the sun through plants and animals, but the total amount of energy in the universe remains the same. The second law states that in an isolated system—such as Earth—entropy, or the incapacity of an energy system to do work, tends to increase over time. The second law is obviously most applicable to human and natural systems constrained by the availability of useable energy in a given environment. Moreover, modern society’s dependence on the concentrated energy of finite fossil fuel reserves makes the second law even more worrisome. For example, when the concentrated energy of petroleum is burned within an internal combustion engine and dissipated into carbon dioxide gas, its capacity to do work is no longer present. As the combustion of fossil fuels yields greater and greater entropy dissipation, the second law of thermodynamics should predict that human societies will inevitably come up against profound and, perhaps, unsolvable energy constraints.

**Cultural Ecology**

During the 1960s—at the same time as the development of systems ecology—many geographers and anthropologists began to apply positivist methods of ecology to the study of human interactions with local environments. A cadre of field-based cultural ecologists began to study human behaviors in terms of functions within larger ecological systems regulating the flows of energy, nutrients, and biomass. Although the original method was to focus on rigorous comparative work to generate specific laws of nature-society interactions along a hierarchy of increasing social organizational complexity (from hunter-gatherers to industrial societies), eventually systems ecology proved to offer more testable quantitative tools of analysis. Energy emerged as a universal common metric with which different scholars could formally engage in cross-cultural comparisons of environmental interactions, a metric that had implications for evaluating the ecological efficiency of cultural practices from a seemingly objective standpoint.

For example, the anthropologist Roy Rappaport offered a systematic examination of the energy flows in the gardening practices among the Maring people of New Guinea. Through intensive field observations, Rappaport was able to estimate the calories expended through the performance of everyday garden maintenance—from clearing the land for tilling to weeding and harvesting. Through these measures, for instance, he was able to estimate how many calories per square feet were needed to clear a particular patch of forest in the New Guinean highlands. He could also estimate the calories made available from the production of potatoes and other vegetables meant for human and pig consumption. Measuring the flows of energy expended in
the production of food against the flows of energy afforded from the food itself allowed the calculation of specific numbers denoting the efficiency of energy expenditures.

Cultural ecologists are notorious for only examining small subsistence-based rural societies such as the Maring people of New Guinea. In the minds of the researchers, such settings seem to provide a model of a “closed” ecological system wherein the flows of energy in and out of the system can be clearly traced and measured. Indeed, such simplicity is a methodological necessity from a holistic systems perspective determined to measure and observe all the relationships constituting a particular system. In the following subsection, we discuss how energy concepts have been applied to the analysis of more complex industrial societies.

**Ecological Economics**

During the energy crisis of the 1970s, it became clear that standard economic production functions based on the contributions of labor, manufactured capital, and materials missed very important variables. Standard economic models ignored, on the one hand, the energy required for nearly every economic action of production and consumption and, on the other hand, the wastes produced through such actions, which themselves often had deleterious ecological effects.

Ecological economics emerged in the 1980s specifically to correct this blind spot in mainstream economic theory. Nicholas Georgescu-Roegen’s landmark work, *The Entropy Law and the Economic Process*, proposed a complete revision of economics to confront the biophysical foundations of the production and consumption of goods and services throughout the economy. Specifically, Georgescu-Roegen made much of the second “entropy law” of thermodynamics to call attention to the inevitable increase in entropy dissipation bound up in fossil fuel–based economies—and, by consequence, the need to move toward a very different energy system. After these pioneering insights, many ecological economists, such as Herman Daly, argued that ecological truisms compel us to move away from growth-based economies toward a “steady-state economy” that would balance overall material and energy throughput with the regenerative capacities of the biosphere.

While early ecological economics was high on theory, it was low in terms of the empirical applications and general testability of cultural ecology. In light of this gap, geographers such as Cutler Cleveland helped develop the groundbreaking concept of “Energy Return on Investment” (EROI) as a means to evaluate empirically the energy efficiency of economic systems. EROI calculates the energy expended in the production of particular energy resources in comparison with the energy obtained. A classic example of the utility of EROI is the case of corn-based ethanol. Although attractive from the perspective of corn growers and lessening the dependence on Middle Eastern oil, EROI reveals that it takes almost as much fossil fuel energy to produce corn-based ethanol as the energy one gets out of it. Cleveland and others have also used EROI to argue that although particular energy resources might be becoming cheaper money-wise—like oil in the late 1980s and 1990s—that doesn’t necessarily mean it takes less energy to produce such oil. EROI can both be used on a microscale in terms of comparing the energetic efficiency of different energy resources (coal, nuclear, solar, wind) for site-specific electricity generation and on the macroeconomic scale, focused on nations as a whole. By showing the economy-wide decline in EROI in the U.S. economy during the 20th century, Cleveland and colleagues were able to demonstrate Georgescu-Roegen’s “entropy law” on a grand scale. A thoroughly ecological EROI analysis could also link metrics on energy expenditures with the actual ecological impacts of energy provisioning. Such a perspective would not only trace the energy expended in certain economic actions (such as the production of energy) but also detail the ecological impacts of economic actions (e.g., water contamination in petroleum production or carbon dioxide produced through electricity generation at coal-fired power plants).

**Critiquing Human Ecological Approaches to Energy**

From a Barrowsian perspective, it is clear that human ecological approaches to energy have made important strides, but they may not be
ecological enough. Whether one is concerned with counting the calories expended in tilling the soil or the amount of energy invested in the production of coal for electricity generation, it is clear that merely tracing the flow of energy does not tell us the entire story of energy-society relationships. A holistic ecological approach must also take into account the political relations shaping access to and control over energy resources and flows. Political ecologists criticize cultural ecologists for approaching agrarian communities as if they were closed systems marked by a functional, homeostatic balance with nature. Indeed, demonstrating that the cultural practices of certain local communities are ecologically functional and energetically efficient does not necessarily mean that they are free of exploitation based on gender, ethnic, and class power relations. Nor is it really possible to treat any community on the planet as a closed ecological system. Nearly all human beings are to some extent embedded in a cash economy, which, by definition, intricately links them into global circuits of capital and commodity flows (and the energy relations therein).

As with cultural ecological examinations of subsistence economies, ecological economists often treat the economy as an isolated conceptual entity. EROI analysis tells us nothing of the dynamic geopolitical relations shaping access to and control over petroleum resources in the Middle East. In energy-intensive societies such as the United States, the consumption of energy resources is also shaped by profound political and cultural narratives situating cheap oil as necessary to a unique “American way of life.” Overall, a relational human approach to energy still has much work to do in terms of tracing the entirety of social, political, and economic relationships that shape energy systems.

Matthew T. Huber

Further Readings


Energy Models

Energy models characterize the energy system, its evolution, and its interactions with the broader economy. The energy system consists of primary resources, including both fossil fuels and renewable sources; electric power plants, refineries, and other technologies to process and convert these resources into secondary energy carriers such as electricity and gasoline; technologies such as furnaces and light-duty vehicles that use these energy carriers to meet demands for energy services; and end-use demands for services such as indoor space heating and personal transportation.

Attention to energy has increased with growing concern over global imbalances between the demand and supply of petroleum resources, as well as the need to mitigate global climate change and other environmental impacts related to energy use. Understanding the role energy plays in these concerns requires an awareness of how energy is produced and consumed, the price dynamics that mediate this relationship, and the physical systems that transform primary energy into secondary energy carriers and ultimately the services energy provides. Modeling frameworks from both economics and operations research provide analytical tools for characterizing the energy system and its evolution and find wide use in design,
ENERGY MODELS

forecasting, scenario analysis, and similar applications. As the energy system is driven by region-specific patterns of resource availability, land use, technology deployment, and demand, energy modeling necessarily requires a geographic perspective. Likewise, spatially referenced energy models are useful for studying emissions related to climate change, infrastructure needs, land use, transportation, and other areas of interest in contemporary geography. This entry provides an introduction to energy models, first reviewing common analytical frameworks before discussing the importance of a spatial perspective.

Energy Modeling Frameworks

Unlike geologic models of petroleum reservoirs, assessments of single technologies, or statistical forecasts of demand, energy models characterize interdependencies between parts of a system or relationships between the energy system and the broader economy. The most commonly used energy models are dynamic, built around a mathematical optimization framework, and incorporate either classical economic theory or systems engineering principles. Outputs include technology choices, energy consumption, pollutant emissions, and marginal costs for energy resources, over time and by region.

Individual energy models differ in their representation of behavior (i.e., what is determined within the model endogenously rather than assumed exogenously), mathematical framework, economic focus, level of technological detail, and sectoral coverage, as well as in their geographic and temporal resolution and horizon. An important modeling distinction captures differences in the representation of economic feedbacks and technology detail. “Top-down” models adopt a macroeconomic perspective and balance the demand for and supply of resources, including energy, across sectors of the economy. The representation of technology in these models, however, tends to be very aggregate. The Second Generation Model provides an example of a top-down model with a substantial energy system focus. “Bottom-up” models, in contrast, downplay key economic feedbacks, such as the responsiveness of demand to changes in energy prices, in favor of a detailed representation of individual technologies.

Partial equilibrium models such as MARKAL, for instance, focus primarily on the physical energy system and determine the least-cost mix of technologies needed to meet specified demands across sectors of the economy. In practice, advances in computational power are blurring the top-down, bottom-up distinction.

Energy models also differ in the interpretation of their results. Models such as MARKAL are optimization based and are therefore prescriptive in orientation. Results reflect what a rational, cost-minimizing decision maker should do, subject to constraints that capture how the energy system functions, as well as policies such as emissions caps that influence the use of resources and technologies. More descriptive frameworks include regional econometric and input-output (IO) models, as well as simulation models. System dynamics models such as the U.S. Department of Energy’s FOSSIL2 and its successor IDEAS (Integrated Dynamic Energy Analysis Simulation), as well as LEAP (Long Range Energy Alternatives Planning System), provide examples of the latter.

Importance of a Spatial Perspective in Energy Modeling

Many energy models are capable of capturing regional differences in resources, transport costs, technology suitability, end-use demand patterns, policies, trading patterns, and other factors that drive energy production and use. Common geographic representations include administrative units such as national or state boundaries and census districts, or energy-system entities such as the North American Electric Reliability Corporation regions and Petroleum Administration for Defense Districts in the United States. Like the representation of time and technology, greater geographic resolution increases model run time, assumes that data are available at finer spatial scales, and compounds the assumptions needed for model calibration.

A spatial perspective, however, remains important for energy modeling. Input parameters, for instance, are often the product of detailed geographic information system (GIS) spatial analyses, even when aggregated up to a broader regional representation for model use. Likewise, GIS modeling provides a means to downscale
and visualize results. Surrogates such as human populations, to give an example, may be used to disaggregate air pollutant emissions from model regions to a finer raster grid. Such applications involve soft coupling of GIS and energy system models. While GIS modeling provides a valuable tool for energy-related analyses, few examples exist of tightly linked GIS energy system models. The AIDAIR modeling framework, which has been used for integrated energy-environmental analyses of Geneva, approaches this ideal.

Timothy Lawrence Johnson

See also Anthropogenic Climate Change; Biofuels; Energy and Human Ecology; Energy Policy; Energy Resources; Geothermal Energy; Solar Energy; Three-Dimensional Data Models; Wind Energy

Further Readings


U.S. Policy

Early federal energy policy involved major subsidies for oil, natural gas, and electric power production; delegation of retail service regulation to the states via the 1935 Public Utility Holding Company Act; and extension of electricity supply to all sectors of society through the Rural Electrification Act of 1936. It also supported new sources, as through the Federal Water Power Act of 1920, licensing nonfederal hydro projects, and the Price-Anderson Act of 1957, providing accident liability limits for nuclear power providers. Subsequent legislation addressed environmental concerns (e.g., the 1970 Clean Air Act and the 1982 Nuclear Waste Policy Act) and energy efficiency (e.g., the 1978 National Energy Conservation Policy Act).

Historically, U.S. energy policy has tended to be event and crisis driven, as with the 1950s “Atoms for Peace” program for nuclear power, the belated creation of the North American Electric Reliability Council to coordinate grid connection following the November 1965 northeast blackout, major changes in nuclear facility licensing following the March 1979 Three Mile Island accident, and President Richard Nixon’s Project Independence to bolster domestic supply and reduce foreign dependence following the 1973 oil

ENERGY POLICY

The procurement and use of energy resources are defining characteristics of human society. Energy policy refers to the myriad collective public and private decisions taken to secure access to and conserve energy resources necessary for sustaining a society, including provision for social reproduction and economic expansion. For the modern nation-state, energy policy extends far beyond the obvious rules and regulations governing public intervention in energy markets. Whether overtly or indirectly, energy policy also involves financial, military, trade, industrial, environmental, transportation, housing, and many other spheres of public policy. While this entry focuses on public, particularly U.S., policy since World War II, a broader recognition of energy policy is suggested.

In the past century, energy policy for the industrialized world has focused on three interrelated areas of continuing concern. These include access to secure resources (e.g., a major rationale for European incursion into the Middle East and Northern Africa), the cost and impact on the economy (especially after the oil “shocks” of the 1970s), and more recently environmental damage and global climate change.
embargo by Arab members of the Organization of the Petroleum Exporting Countries (OPEC). While most of the 1970s policy initiatives failed, one notable success was approval of Corporate Average Fuel Economy (CAFE) standards for cars and light trucks under the Energy Policy and Conservation Act of 1975, though these were not raised beyond 1985 levels until 2007.

With just 4.5% of the world’s population, the United States consumes a quarter of the world’s nonrenewable fossil fuel energy and, by extension, contributes a similar amount to global greenhouse gas (GHG) emissions. Energy-generated carbon dioxide alone is responsible for 82% of total U.S. GHG emissions, principally through electric generation (41%) and transportation (33%). An abundant domestic supply allows coal to be used for over half of all electricity production. However, with combustion causing air pollution, coal policy since the 1970s has focused on a reduction in particulates and in nitrogen oxide and sulfur dioxide emissions (the latter causing acid rain). Related acknowledgment of coal’s major contribution to global warming led to federal support during the George W. Bush administration for carbon dioxide capture and geologic storage. Modeled after the successful sulfur dioxide trading program under the 1990 Clean Air Act Amendments, a market-based approach to carbon reduction through emission cap and trade undertaken by the European Union (EU) to meet its obligations under the Kyoto Protocol is slowly gaining political acceptance in the United States.

With training in nuclear engineering, President Jimmy Carter provided the boldest energy program of any recent administration. This culminated in the creation of the U.S. Department of Energy (DOE) in 1977—reorganized from more than 50 different federal agencies and with responsibility for a biennial National Energy Policy Plan—and the creation of the independent Federal Energy Regulatory Commission (FERC) from the preexisting Federal Power Commission. The FERC regulates the interstate transmission of natural gas, oil, and electricity, as well as natural gas and hydropower projects. From the late 1970s through the 1990s, federal policy shifted to remove regulations on oil and gas prices and encouraged states to restructure preexisting monopoly gas and electric utilities through provision of grid access for independent generators, while allowing consumer choice with supply.

On taking office in 1981, President Ronald Reagan cut most federal spending for renewable energy research and allowed federal tax credits for production and residential installation to expire. Under his leadership, and generally continuing through the Bill Clinton and both Bush administrations, federal policy addressed increasing domestic fossil fuel supply through tax breaks, subsidies, price deregulation, and supply restrictions while avoiding demand management, such as through increased CAFE mileage standards or significant attention to renewable energy. The common belief was that liberalized markets would bring retail costs down and encourage new investment. During this period, new wind and solar supply tended to rise and fall according to the availability of a modest (compared with fossil fuel and nuclear power support) federal production tax credit. Vacillations in federal and state support notwithstanding, actual investment in wind and solar energy continued to grow due to a rapid increase in fossil fuel prices and utility concern with price stability. Proponents of renewable energy argue that it could compete on its own should the subsidized playing field be leveled.

During the Reagan administration and early in the George H. W. Bush administration, as global petroleum prices fell, U.S. energy policy continued to be reactive, with scant attention given to either conservation or the renewable energy supply. In place, neoliberal energy policy focused on deregulation of preexisting monopolies and privatization of public supply wherever possible. However, the affirmation that continued access to the vast Middle East petroleum reserves was of vital national strategic interest—a view first articulated by President Franklin Roosevelt and later embraced by Carter—continued to resonate with subsequent administrations, eventually reaching a crescendo under the George W. Bush administration. Thus, the United States provided guarantees of military support for accommodating regimes such as that of the Saudi royal family. It also employed covert or direct military intervention when faced with perceived threats to global trade in petroleum, as occurred in Iran during the 1950s and more recently with Iraq. In this manner, U.S. energy policy informs foreign relations and military policy.
Driven by environmental and energy security concerns, provision of renewable energy is now a major policy issue. Half the U.S. states have mandatory renewable portfolio standards (RPS), whereby a set percentage of utility output must come from certified sources. Adopted in the absence of federal standards and favored by free marketers, the state RPS is an alternative to the feed-in tariffs (FIT) deployed by Germany, Denmark, Spain, and most recently Australia and California. Here, utilities must purchase renewable energy at the geographic and technological cost of generation. This allows for a more secure investment climate for generators and leads to deployment over a wider geographic region with greater renewable energy penetration.

The most controversial national energy plan in the United States was prepared in 2001 by a panel chaired by Vice President Dick Cheney with representation from the major oil and energy trading firms (e.g., Enron) but with no public or environmental interest groups invited. With court challenges ongoing to release panel records, the National Energy Policy Report continued to serve as the basis for the Bush administration’s energy policy. Described as a “strategy of maximum extraction” by the energy and defense correspondent Michael Klare, the Bush-Cheney plan blamed the rising prices on the lack of resolve to open up domestic supply, such as in the Arctic National Wildlife Refuge. With scant attention to demand management, the report proposed meeting the projected 1.8% annual growth in electric consumption with the installation of 1,300 to 1,900 new 300-MW (megawatt) (average) power plants through 2020.

As of late 2009, few of these proposals had been approved, with private investment and utility commitment still hampered by very high construction costs despite improved federal support in recent years. However, advocates anticipate that prospects for nuclear power will improve should the nation adopt a cap-and-trade program or carbon tax for GHG emissions. In the meantime, since 2008, major advances in hydrofracturing of deep shale formations contributed to a 35% increase in proven domestic natural gas reserves, with the “bridge” fuel to lower carbon emissions once again supporting new electric power construction.

Congress and the Bush administration supported development of hydrogen and ethanol as alternative transportation fuels through the Energy Policy Act of 2005 (with a goal of 100,000 hydrogen-fueled vehicles by 2010 and tax subsidies for biofuels and “advanced” nuclear reactors) and the Energy Independence and Security Act of 2007 (modestly raising CAFE standards to 35 miles per gallon by 2020 along with a continued subsidy and a major mandate for 36 billion gallons of biofuels by 2022). Critics were quick to point out, however, that these alternative fuels were only as “green” as the resources used to prepare them, with hydrogen stripped from fossil fuels or released via water hydrolysis, while corn-based ethanol is also linked to fertilizer runoff and rising world grain prices. By 2008, these concerns and international food riots led to a reassessment of hydrogen and biofuel needs in Europe, which had adopted similar goals, while the U.S. program has since been cut back under the Obama administration.

While neoliberal proponents of deregulation claimed that market competition would reduce retail prices and promote the independent renewable energy supply, the record with electricity is mixed, with no discernable trend toward either lower prices or increased renewable energy in the states undertaking market reforms compared with those remaining with regulated monopolies. The California electricity “crisis” of 1999–2001, culminating in widespread power outages, underscored what can happen when poorly devised market reform encourages independent traders, such as Enron, to “game” the system. In the interim, further restructuring is on hold in the United States. On the other hand, most EU nations have succumbed to the lure of market reform, with national monopolies broken up. Here, initial success with renewable energy provision may be attributed to national, and now unified EU, reform, rather than reliance on piecemeal state implementation as in the United States.

Energy Policy in Other Countries

Energy policy in most nations reflects the availability of domestic resources as well as foreign relations and international trade. Japan, for instance, has virtually no petroleum, so while its
policy has promoted nuclear power, solar energy, and energy efficiency, it has not been able to avoid a significant oil import dependence. Western Europe has similarly supported energy efficiency but has also used high taxes to control energy demand (as has Japan). European nations such as Germany, the United Kingdom, and Poland used to be major coal producers and consumers, though with diminished resource quality these nations have increasingly turned to other sources as they have phased out subsidies and reformed their economies. The United Kingdom capitalized on its North Sea oil reserves, which provided a temporary respite to foreign dependence in the 1970s through 1990s, though since then it has had to look elsewhere. Even Norway’s vast North Sea petroleum reserves have been dwindling more recently. Energy policy in the EU since the late 1990s has focused on GHG reduction, with wind power development in Germany, Spain, and Denmark being notable success stories thanks to the FITs. Scandinavian countries, for their part, continue to promote the use of their substantial hydroelectric resources and geothermal energy (Iceland only). Another key regional player is Russia. Having revived its struggling fossil fuels industries following post-1990s economic restructuring, it is by far the largest fossil fuel producer in Europe.

In recent years, the rising electric and petroleum demand has been greatest in developing nations, such as India and China, the latter now responsible for a quarter of global greenhouse gas emissions. Even so, coal use is booming and supported by government policy, especially in China, the largest coal producer in the world. As is common in Russia and among OPEC nations, various subsidies contribute to increased domestic consumption, even as China and India are also growing leaders with wind and solar development. China is also the world’s leading proponent of expanded use of nuclear power. In 2007 and 2008, China and India, respectively, published national climate change action plans emphasizing energy efficiency and renewable energy—although not at the cost of economic development. In the case of China, its automobile mileage efficiency standards now exceed those of the United States, which has been driven by the need to control foreign oil dependence, while China’s domestic oil industry continues to develop.

Divergent discourse notwithstanding, the reality of global warming and the coming peak in global oil production, with much higher fossil fuel energy prices, will continue to push the United States, along with the rest of the industrialized and developing world, toward a common understanding of the need for a forceful reduction in demand and major expansion of sustainable supply. The required economic transformation and social accommodations will be revolutionary, but consensus has grown that such changes must be made. While national needs, however defined, are likely to continue to be the basis for most energy policy, an increasing number of international forums and organizations have been created to further common interests.

Michael K. Heiman and Barry D. Solomon

See also Carbon Trading and Carbon Offsets; Climate Policy; Energy and Human Ecology; Energy Models; Energy Resources; Geothermal Energy; Hydroelectric Power; Market-Based Environmental Regulation; Neoliberal Environmental Policy; Nuclear Energy; Public Policy, Geography of; Solar Energy; Wind Energy

Further Readings

A distinction is made in the energy field between energy resources and energy reserves. Resources refer to the proportion of the total resource base that has been identified with at least a small degree of certainty and that might be recovered in the future. Energy reserves, in contrast, indicate the proportion of the total resource base that has been identified with certainty and that can be economically recovered and used with current technology. Both quantities are spatially specific and not fixed. However, estimated quantities of renewable sources of energy are highly variable and uncertain, given the highly uneven geographic distribution of resources and technology to capture sources such as solar, wind, and biomass energies.

It is useful to know the geographic distribution of energy reserves as well as production, consumption, and trade patterns of energy commodities such as petroleum and coal as these figures are used to calculate the energy import dependence and scarcity of fuels in a nation, region, or the world. Energy resource data, while they are also potentially useful, are updated less frequently and are much more speculative than are reserves. Extensive nonrenewable energy resource data at a variety of geographic scales are available from organizations such as the U.S. Department of Energy, International Energy Agency, the United Nations, and private corporations or trade associations such as British Petroleum (BP), the European Wind Energy Association, or the Solar Energy Industries Association. Helpful statistics include the energy resource-to-production ratio for oil or gas in a country; national or state rankings of energy resources, reserves, and production for oil, gas, coal, and uranium (Figure 1); and wind energy resource by state or geothermal energy production by nation. It is widely recognized that petroleum and other fossil fuels are finite resources that contribute to local pollution problems and global climate change and instability, and for these and other reasons, renewable sources of energy will be more rapidly developed and deployed in the future. Despite this, geopolitical and global economic and military affairs are still intertwined with the use of fossil fuels, especially petroleum products.

**Figure 1** Global reserves of nonrenewable energy sources by type

Fossil fuels are divided into conventional sources such as crude oil, natural gas, and coal, and unconventional sources such as tar sands, shale oils, and methane hydrates (clathrates). Their commonality is that they are nonrenewable hydrocarbons found within the top layer of the Earth’s crust or in the ocean floor and were formed over millions of years from the remains of dead animals and plants. The top five nations today in terms of oil reserves are in the Middle East, with the United States a distant 10th. Perhaps surprisingly, the Middle East’s share of global proved oil reserves has changed little in the past half century, accounting for around 60% (Figure 2).

Much of global oil production is dominated by the 12 member nations of the Organization for the Petroleum Exporting Countries (OPEC). Half of its members are in the Middle East, including Saudi Arabia, the world’s largest producer. Once accounting for over half of global production, OPEC’s market share has varied through time and today stands at around 45%. Other leaders in crude oil production include Russia (second, at 12.3%), the United States (third, at 8%), China (fifth, at 4.7%), and Mexico (sixth, at 4.7%). In recent years, there has been much debate about when global oil production will peak or plateau (U.S. oil production peaked in 1970), with some arguing that it has already occurred, while others claim that the growth can continue unabated for 20 to 30 years.

The consumption of refined petroleum products is concentrated in North America and Europe. Europe/Eurasia and the United States account for almost half of the global refining capacity, or about 20% and 29%, respectively. With the exception of Russia, many of the countries in these regions, along with Japan and China, are thus heavily dependent on oil imports (crude oil and refined products); for example, the United States is around 60% dependent on net petroleum imports. Most U.S. imports come from Canada, Saudi Arabia, and Mexico.

Synthetic forms of oil supplies could technically replace petroleum, and the most market-ready
option is tar sands (also called bitumen or petroleum sources with high viscosity). The vast majority of these resources have been produced in Northern Alberta, Canada (in the Athabasca region), and Venezuela (in the Orinoco region in the eastern part of the country). Canada has been the leading producer as it has more reserves of tar sands than conventional oil. The United States has even larger reserves and more than 60% of the world’s resources (found mostly in Western states such as Utah and Alaska), but it has yet to develop them. Another oil alternative is oil shale (also called kerogen), which is found in fine-grade sedimentary rocks in the Western United States and many other countries. The kerogen can be converted into synthetic crude oil via pyrolysis. Oil shale has been produced on a very limited scale in more than a half dozen countries. U.S. reserves are similar to those of its oil supply, but again commercial production has yet to commence. The Federal Bureau of Land Management has leased five research and development tracks to oil companies in Western Colorado. Production could occur after 2010 at the earliest, probably using in situ processing. Barriers to the development of these unconventional energy resources are mostly economic and environmental, with water pollution, waste generation, greenhouse gas emissions, and water requirements potentially huge. In addition, both oil shale and tar sands have much lower volumetric energy density than conventional crude oil.

Natural gas resources are found together with oil (so-called associated gas) as well as independently. While gas (which is 80–95% methane) is cleaner than oil, it was less intensely explored and developed until the 1950s, is inconvenient to store, and is much more costly to ship because of its low energy density. Overseas shipment of natural gas requires that it be first liquefied to 1/600 of its normal volume and at cryogenic temperatures. The procedure involved is costly. The gas reserve base is more geographically concentrated than is oil, with 55% of the world total found in Russia, Iran, and Qatar (Figure 3). Russia and the United States dominate global gas production, accounting for around 40% of the total, with

![Figure 3](image-url)
more than half of the gas consumption taking place in the United States and in countries of the Soviet Union and the European Union. The date of peak global gas production is even more uncertain than for peak oil, though U.S. production peaked in 1973.

Several unconventional gas sources could be exploited. The most important of these are coal-bed methane from underground coal mines and methane hydrates (clathrates). The latter are crystalline solids containing gas molecules surrounded by ice, which are found in outer continental shelf sediments and in polar regions. While methane hydrates resources are vast, they are yet to be developed. The United States has coal-bed methane reserves equal to about half those of conventional gas, which already account for 8% of total gas production.

The last of the three conventional fossil fuels, and the one by far in greatest abundance, is coal. It is also the easiest one to find and develop, and consequently, it fueled the Industrial Revolution, first in Europe and later in the United States. Global coal reserves are even more geographically concentrated than those of natural gas, with the United States, Russia, and China accounting for 61% of the total; Australia, India, and South Africa account for an additional 21% (Figure 4). China is by far the largest producer, accounting for more than 41% of global production in 2007, more than double the production of the United States, at number two. Despite the growing global concern over controlling carbon dioxide emissions, only Europe (except for Russia) and North America, among the major producers, have experienced a leveling of production over the past decade. Coal consumption patterns are somewhat similar to those of production, with Asia dominating (60% in 2007), followed by North America and Europe. Major coal exporters include Australia, Indonesia,
Russia, South Africa, and Colombia, while major coal importers with little domestic production include Japan and South Korea.

Coal resources are primarily used as solid fuel for electric power generation as well as for some heating; it has a secondary market as coking coal for use as a fuel and reducing agent for smelting iron ore in blast furnaces for steelmaking. There is also increasing interest in the use of coal for technologies such as gasification (including methanation, or conversion into natural gas) and liquefaction, with increasing commercial development around the world.

### Biomass Energy

Biomass provides the most important source of renewable energy in the world, but it is renewable only if it is used sustainably. Firewood and charcoal were among the initial energy sources of traditional societies, along with the sun, wind, water, human and animal muscle, crop residues, dried dung, and even gunpowder. Historical evidence shows that wood in any available form and other combustible resources were the first inanimate sources of energy for humanity. Other forms of biomass energy are readily available besides wood, such as dung, animal fats and tallow, agricultural crops and residues, and grasses. The use and benefits of biomass energy resources should be balanced against their other purposes, including agriculture and other nonenergy applications, as well as the many other sources of energy against which they compete. Consequently, statistics on the energy resource or reserve base for biomass and other renewable sources may not be very meaningful.

Much of the work of geographers on biomass energy has focused on the prominent role of fuelwood for home heating and cooking in developing countries, especially sub-Saharan Africa, India, Pakistan, and China. In many of these nations, biomass energy use accounts for 90% or more of total energy consumption, while the comparable figure for the whole world is 10% to 12%. These uses of biomass are often referred to as traditional or first generation, and most of this biomass energy is not traded in markets. Particular concerns of geographers include understanding the socioeconomic and cultural factors that influence deforestation, sustainability, and health problems from inhalation of fine particles from cook stoves. While biomass energy accounts for just 3.6% today in the United States, it was the major energy source of the nation until 1885, when it was overtaken by coal.

Modern or second-generation sources of biomass include feedstock for alcohol fuels, such as sugarcane ethanol in Brazil, corn ethanol in the United States, and biodiesel from rapeseed in Germany or soybeans in the United States. Ethanol can also be made from wood, as can methanol (wood methanol was popular in the northeastern United States from the 1890s until the 1930s, when it was replaced by oil). Advanced ethanol based on cellulosic feedstock such as agricultural and forestry residues, switchgrass, and municipal solid wastes is in the process of being commercialized in several nations, but it will take a decade or more to establish significant market shares.

### Hydropower

After biomass energy, the next largest source of renewable energy in the world is hydropower, which uses the force of moving water to generate electricity or do mechanical work. It is difficult to account for hydropower reserves because of the many competing uses for water. In addition, while fresh water is clearly a renewable source of energy, the lifetime of hydroelectric dams is quite variable because of eventual siltation of the reservoirs and variable rainfall and snowmelt. While most hydropower is used for electricity generation, it has also been used for irrigation and to operate machinery. In many parts of the world, hydropower is the cheapest source of renewable electricity.

The leading producers of hydroelectricity include China, Canada, Brazil, the United States, and Russia. The United States used to be the world’s largest generator, but its output has been declining. Among the major producers of hydroelectric power, Canada and Brazil are the most reliant on it as water accounts for 65% to 85% of their respective electricity needs, depending on the year and weather conditions. Among some smaller countries, reliance on hydroelectric power is even larger, as it accounts for 95% to 99% of power needs in Scandinavian countries such as Norway and Iceland.
Since the best dam sites have long been tapped for hydroelectricity in the United States (and some of these sites have been decommissioned in recent years), most of the major new projects are being built in the developing countries. These include China, India, Malaysia, Vietnam, Brazil, and Azerbaijan. The most famous of these is the 22,500-MW Three Gorges Dam on China’s Yangtze River, the world’s largest hydroelectric plant.

Other Sources of Renewable Energy

There are several other sources of renewable energy. One of these, wind energy, has been the fastest-growing energy source in the world since the early 1990s. Available power is a function of the wind velocity cubed, so good steady winds of 15 miles per hour or more are especially desirable. However, intermittency is a problem with wind (as is the case with several renewable energy sources). Germany is by far the leading user of wind power in the world, followed by the United States and Spain, while India and China are the leading users among developing countries. Denmark has the highest rate of per capita reliance on wind for electricity, at around 15%. The United States, once the leading wind energy user, has had fluctuating demand growth in the past decade. The state of Texas is now the leading wind user in the nation, followed by California.

Solar energy, the demand for which is also growing rapidly, has a much smaller market. Western nations such as Germany, Spain, the United States, and Japan have been working to increase the conversion efficiency of photovoltaic cells but mostly sell space and water heaters. The cost of photovoltaics continues to fall, but they are competitive only in limited applications. Also, more than 10 solar power plants using parabolic troughs have been built in the Southwestern United States. Developing nations, in contrast, are much more interested in low-cost solar cookers, especially in rural areas.

Geothermal energy, although site specific and depletable on local scales, is generally still considered renewable. The United States and the Philippines are the largest users of geothermal electricity, though the United States has not experienced capacity growth in the past decade. Besides power production, geothermal energy can also be used for space heating, aquaculture, and heating of swimming pools, spas, resorts, and greenhouses.

Finally, some remaining renewable energy sources that have not been seriously developed include tidal power, wave energy, and ocean thermal electric conversion.

Uranium and Nuclear Power

Uranium is used for fuel to generate electricity in nuclear power plants, which were first commercialized in the United Kingdom, Russia, and the United States in the 1950s. This naturally occurring and widely distributed element is found in minerals, soils, and seawater around the world. Australia, Kazakhstan, and Canada account for 50% of the world’s uranium reserves, with Canada, the United States and South Africa accounting for another 23% (Figure 5). Naturally occurring uranium contains more than 99% uranium-238 (238U) and about 0.7% uranium-235 (235U). To provide fuel for most nuclear power plants, the 235U concentration must be enriched to around 3%. Since fuel is only a small portion of the operating expense at nuclear plants (the majority is the capital investment), electric utilities in major countries such as the United States will import a majority or all of their uranium fuel for cost reasons even if there is a large domestic resource base.

In rank order, the leading generators of nuclear power in the world are the United States, France, Japan, Russia, and Germany. The most nuclear-dependent countries are France and Lithuania, which generate 77% and 64% of their electricity from nuclear power.

End Use Efficiency and Conservation as Energy Resources

While end-use efficiency and energy conservation are not a traditional energy “resource,” they can provide low-cost energy services in all countries. Although the technical, economical, and social potential of such technologies varies by sector and country, these options are alternatives that can be considered to meet a large share of energy needs. There is no consensus as to how large this
potential is in most countries. Consequently, energy audits may be very useful to determine the potential for these technologies in buildings and in industrial sectors, for example, through the use of more efficient refrigerators and other appliances, lighting, ballasts, and motors and cogeneration of electricity and heat from currently discarded industrial process heat. Attractive transportation options include greater use of public transit, telecommuting, hybrid gasoline-electric vehicles, plug-in hybrids, and diesel-fueled vehicles. Most of these options are readily available. Many nations have government-sponsored energy-efficiency programs and centers, which often provide grants, loans, and technical assistance to meet the potential for energy-efficiency improvements. Industrial groups, such as the U.S. Electric Power Research Institute, provide similar services.

Barry D. Solomon

See also Coal; Energy and Human Ecology; Energy Models; Energy Policy; Geothermal Energy; Hydroelectric Power; Market-Based Environmental Regulation; Nonrenewable Resources; Nuclear Energy; Petroleum; Political Economy of Resources; Renewable Resources; Solar Energy; Wind Energy

Further Readings


The Enlightenment has typically been defined as a period of European thought, a broad intellectual movement, or the Age of Reason that took place in Europe approximately from the years 1685 to 1815 and that emphasized experience, reason, and mistrust of religion and traditional authority.

Definitions such as this one seem to confine the Enlightenment to a universal, static, and singular movement that existed within specific European countries. Such definitions also imply that it was understood and expressed similarly in all countries despite differences in languages, politics, economics, and religious beliefs. Nevertheless, the Enlightenment as a discourse and as a process was far more complex than what was originally envisioned, reaching more than one continent and being constructed in discrete and multiple forms around the globe. Scholars should not speak of the Enlightenment as a unitary phenomenon but rather as a process that was manifested in a variety of ways pending its locus of enunciation and reception. In this sense, the Enlightenment should be perceived, as Charles Withers notes, as a dynamic phenomenon characterized by a variety of local and temporal factors. Following Withers’s suggestion, we should then think about the Enlightenment in the plural—the Enlightenments—understanding it as a dynamic transnational phenomenon. The Enlightenment ought to be considered as a series of processes, discourses, and lines of thought manifested in multiple ways in areas as diverse as religion, politics, science, literature, art, education, economy, and philosophy, to name a few.

The Enlightenment also implied a time in which the production of knowledge traveled at a global scale through printed books, maps, scientific expeditions, and economic trading. Political, economic, religious, and cultural transactions reflected many of the main premises of the Enlightenment with their emphasis on critical reason, progress, utility, and order. Within this context, space constituted an important element in the discourses produced during this time. Not only the locus of enunciation determined the nature of the Enlightenment as a process, but the space where it was read represented in itself another relevant factor when understanding the multiple layers of the Enlightenment. The famous ideological precepts of the Enlightenment, such as order, utility, progress, and reason, were articulated in different ways in countries such as France, Britain, Spain, Mexico, and Peru, to name a few. To this extent, as Livingstone and Withers suggest, the European Enlightenment was substantially molded through engagement with other parts of the world, notably with the Americas as a whole and with the Pacific. Furthermore, the Americas as well as the Pacific produced their own Enlightenments. To this extent, we can envision the Enlightenment as geographical in nature.

Livingstone and Withers have contributed greatly to our understanding of the Enlightenment as a geographical phenomenon. They both call attention to the centrality of space in the discourses belonging to this period. For them, geography encompasses the sites and practices in which the Enlightenment as a process took place. Also, according to them, the Enlightenment as a series of ideas was produced, debated, and consumed in multiple places and spaces. Exploring, traveling, mapping, and naming became key activities in such processes. Nature itself became a commodity, while traveling, mapping, and naming served as critical vehicles to inquire, survey, categorize, and give order to the world.

The obsession of Enlightenment intellectuals to study, categorize, and give meaning to the natural and physical world rendered geography, as it was practiced at the time, a perfect tool to understand and explain the world. Geography did not exist as a discipline in the 18th century. It was instead considered part of the empirical sciences, in which experience and the senses were
two crucial elements to understand the world. As Withers defines it, in the 18th century, geography constituted the descriptive study of the Earth as a whole, either in textual or in mathematical terms. However, the study of the Earth was a subject that guided the intellectual and political endeavors of philosophers, theologians, cosmographers, mathematicians, and even merchants. The local, or faraway, countries, kingdoms, and continents became the focus of their taxonomic studies as well as their political, religious, and economic interests. To understand the world and its spatial nature implied knowledge of it and consequently the ability to order it and master it. For Enlightenment thinkers, to achieve human progress, it was necessary to know the potential of the world they lived in. As Withers suggests, to think critically “in and about the world” became a significant trait of the age of the Enlightenment. For the intellectuals, that critical thinking involved the act of positioning themselves into or within a privilege location that would deem them wholly capable to describe and explain the world. That position constituted that ideal point of observation where the observer felt that he could explain the world in a transparent fashion or in its pure factual nature, free of any particular moral, political, or religious beliefs. The accumulation of geographical knowledge during the culture of the Enlightenment therefore became a crucial instrument to legitimate, reproduce, and perpetuate a specific social order.

The Enlightenment constituted a key element in understanding the origins and the nature of the world. Humans, plants, and animals were observed, studied, and explained in terms of where they were located. Place as well as space gave new meaning to the human and physical world. In this sense, the Enlightenment had its own geographies as it took place in and over space.

To understand the complexities and ambivalences that characterized the philosophies of the Enlightenment in its search for reason, order, and progress, one must recognize, on the one hand, its multiple places of enunciation and, on the other, the different manners in which it was received and read around the world. The Enlightenment required understanding beyond the limits of Europe as it implied a transoceanic movement that was transformed depending on its multiple locations. The dynamic sites of production as well as reception constitute elements of the Enlightenment that call attention to the Enlightenment as a process, as a thought, and as a movement that challenges any scholar who approaches it as a singular cultural, scientific, and political form.

Mariselle Meléndez

See also Cultural Geography; Human Geography, History of; Livingstone, David

Further Readings

Enterprise GIS

Enterprise GIS refers to geographical information systems (GIS) in private companies, governmental organizations, or any similar entity. The main requirement here is that the company or organization consists of a number of departments or other areas with professional activities that require at least some support of a GIS.

Requirements

In the early GIS years, GIS users were mainly working at separated and sometimes isolated workstations. Today, workstations are connected, linked to other applications, and therefore integrated in the whole information technology (IT) infrastructure of an organization.
The requirements for an enterprise GIS are as follows:

- Easy and flexible modeling of the business logics of the different processes
- Support of the different functional requirements in the departments from road construction to planning and facility management, just to mention a few
- Support of different user groups, experienced as well as occasional users
- Efficient organization of data commonly used internally by individuals or departments
- Ability to exchange data, integrate, or link data and software components of different GIS and other applications

**Technical Issues**

To establish an enterprise GIS, a number of IT-techniques/technologies have to be considered. The most important ones are as follows:

- Client-server architecture (see Figure 1)
- (Distributed) database managements systems allowing multiuser access, definition of access rights, and so on—sometimes also data warehouses
- Sophisticated data management and integration tools
- IT standards such as SOAP or WSDL to implement service-oriented architectures (SOA)
- GI standards such as GML and OGC Web services such as WMS and WFS
- Work flow support

GIS vendors offer specific modules to build an enterprise GIS that in general support the techniques listed here.

**Organizational Issues**

Establishing an enterprise GIS requires complex technologies, but the real challenge is often an organizational one. All the potential categories of
users, their processes, the different kinds of data, functional requirements and software products, and the systems that must be linked have to be known, and a corresponding system architecture has to be defined. It is often a balance between having an IT landscape that is homogeneous and satisfying the different requirements of the various user groups.

Figure 2 gives an overview of one example of the organization of an enterprise GIS. Technically, it is client-server based and includes one central database and different GIS clients with external databases in the departments. Each of them can offer GIS-based services for the whole enterprise; however, many organizational decisions have to be made (who is doing what, where, in which way, and so on).

The organization of an enterprise GIS varies depending on the size of the enterprise, the complexity of the processes, the form and number of integrations of other systems, and other aspects mentioned above. One common approach for an enterprise GIS does not exist.

Wolfgang Peter Reinhardt

See also Business Models for Geographic Information Systems; Client-Server Architecture; Dangermond, Jack; GIS in Local Government; GIS in Public Policy; GIS in Urban Planning; GIS in Utilities; GIS Software

Further Readings


Environmental certification is a procedure by which a third or second party gives written assurance that a product, process, or service is in conformity with certain environmental standards. It is used to give environmentally conscious consumers assurances on environmental aspects of the product or service itself or the production process.

**Process**

The producer (first party) or the buyer (second party) applies for certification to a third party that has created or adheres to a set of standards. **Product standards** are specifications and criteria characterizing products, while process standards characterize the way products are made or services are delivered. **Process standards** can be further subdivided into management system standards, which set criteria for management procedures such as monitoring and documenting, and performance standards, which define what actually happens, for example, whether certain pesticides and fertilizers are used. Because conditions vary worldwide, generic standards are often used, which provide a framework for a more local standard setting. Products or processes are checked against these standards in inspections, which are carried out either by the standard-setting certification bodies themselves or by an inspection body subcontracted by the certification body. The certification decision is then based on the inspection report and possibly other information provided by the producer. Successful certification is documented by a certificate, which can be shown to a buyer and is often signified by a label on products to show to final consumers. Standard-setting organizations for environmental standards can be governments, industry or non-governmental organizations (NGOs), or mixed stakeholder partnerships between these actors.

**Origins**

Environmental certification has existed for over 30 yrs. (years). In 1978, the German government sponsored a standards-backed environmental label, “Blauer Engel” (blue angel). A multistakeholder jury awards the label to products that are more environmentally friendly than are competitor products, such as chlorofluorocarbon-free spray cans, recycled paper, and environmentally friendly heating systems. In 2008, around 10,000 products and services carried the label.

At an international level, the Agenda 21, a global plan of action for sustainable development adopted at the 1992 United Nations Conference on Environment and Development in Rio de Janeiro, called for the introduction of environmental certification and labeling schemes.

**Example 1: Organic Agriculture**

Today, one of the most well-known environmental standards focuses on “organic” production. From the 1920s onward, farmers were inspired by the ideas of Rudolf Steiner and others to develop production methods, which emphasized biological processes and minimized the use of nonrenewable inputs. They began to codify their methods and formed the International Federation of Organic Agriculture Movements (IFOAM), which formulated basic standards in 1980. These standards, which are reviewed every 2 yrs., serve as a guideline on which subsequent national regulation has been based. Organic standards, devised by governments and various IFOAM-accredited NGOs, now exist for all crops and some livestock. Organic standards include various process standards prohibiting, for example, the use of synthetic fertilizers and pesticides.

There are several other environmental certification schemes, including the ISO 14001 Standard for Environmental Management Systems and the Rainforest Alliance certification scheme, focusing originally on habitat conservation. Environmental certification schemes are often linked to public discourses around environmental concerns. They are linked, for example, to dolphin-friendly fishing methods, 100% recycled paper, non–genetically modified ingredients, non-air-freighted fruit, and vegetables and energy-efficient white goods.

**Example 2: Forest Certification**

The 1992 Rio Conference agreed on Forest Principles. Worldwide forest certification schemes,
such as the NGO Forest Stewardship Council (FSC), developed general international principles and criteria and accredited organizations, which in turn certify sustainably managed forest land. In 2008, 100 million hectares of forest in 79 countries was FSC certified. Other schemes include the Programme for Endorsement of Forest Certification, the UK Woodlands Assurance Scheme, the Sustainable Forestry Initiative, and Canada’s National Standard on Sustainable Forest Management CAN/CSA-Z809. There have been several moves to harmonize these different schemes, but so far, these have not been successful.

Current Debates

Currently, several issues related to environmental certification are the subject of debates. These include integration of social and environmental certification, balancing global reach and local specificity, and accountability of those who set the standards.

Link between social and environmental certification. There is a need to further integrate social certification schemes such as Fair Trade, which mention environmental concerns, and environmental certification schemes such as organic certification, which include references to social issues.

Global reach and local specificity of standards. As producer circumstances, climates, soils, and land tenure systems vary, matching worldwide standards to the local reality is challenging. The aim is to keep standards focused on objectively verifiable performance goals (e.g., water quality) while not being overly prescriptive to producers.

Accountability of standard-setting NGOs. Certification and labeling are skewed toward consumers who have the economic means to choose more “ethical” products. Furthermore, NGOs acting as standard-setting or certification bodies are not systematically accountable to producers or consumers, although many invite these stakeholders onto their councils and advisory boards. Also, commercial auditing and certification have increased the risk that inspections can be superficial, affected by conflict of interest, or too focused on details of the standard, without adequate recognition of the producers’ situation. Producers generally pay the inspection and compliance costs, and in turn, they receive a better price or marketing advantage for products or services. Other actors in the chain may benefit without having paid the cost of compliance or certification.

Future Market Developments

As the demand for environmentally certified products increases, large retailers are offering such products. Their requirements for low administrative costs, uniformity, and traceability favor larger-scale producers, putting smaller producers at risk of being marginalized. Another important question is whether consumers can be persuaded to adopt a nuanced view of the different levels of environmental certification. Another key development for environmental certification is the developing global market for carbon emissions trading, where businesses or governments buy carbon certificates from certified carbon reduction and carbon capture projects to offset their own emissions.

Dorothea Kleine

See also Carbon Trading and Carbon Offsets; Fair Trade and Environmental Certification; Organic Agriculture; Sustainable Development; Sustainable Forestry; Sustainable Production; United Nations Conference on Environment and Development

Further Readings


Environmental determinism, also referred to as geographic determinism or even environmentalism, is the contention that the environment holds sway over humanity, controlling human actions and exercising a dominant influence on human development. While a range of perspectives exists on the degree of causal determinism that should be attributed to the environment, advocates of environmental determinism are unified in their recognition that the environment has an overriding influence on humanity. The implication is that the social and cultural behavior of humans is determined to a large extent by environmental factors that are asserted to exist beyond the influence of humanity. This conception of the relationship between humans and the environment has roots that stretch deep into classical history, possibly representing the earliest attempts to explain variations in sociocultural traits across geographic regions. Its compelling simplicity and promise of causal explanations for complex social phenomena have driven recurrent interest in environmentally deterministic theories up to the present.

An Ancient Belief

Belief in the power of the environment to shape society and mold human characteristics has been widely espoused since ancient times. Hippocrates, in his work On Airs, Waters and Places, asserts a connection between world climates and the perceived characteristics of different cultures. The proclivity of Northern Europeans to being poverty-stricken could be attributed to their poor environment, while more climatically favorable regions led to easygoing and amiable societies. Plato and Aristotle believed that the influence of climate extended to the development of governments and was ultimately responsible for stimulating the progress of societies. An acceptance of this direction of influence, strictly from the environment to humans, persisted throughout many centuries and found renewed interest during the Enlightenment. Montesquieu contended that the climate was a potent influence on human behavior, asserting that cold climates led people to be stronger in spirit and body, more trustful and straightforward, and less cunning, while the heat of the lower latitudes resulted in passivity and indolence. As during ancient times, this profound influence on human behavior was believed to be the fundamental explanation for varying forms of government and social organization. For example, it was believed that monarchies tended to occur in more fertile environments, while republics prevailed where the climate was harsh and people were pressured to better their circumstances. Despite the claim that the environment had a vast influence on human life, the environment was rarely the subject of study by environmental determinists as it was commonly viewed as an inexplicable force and considered important only in accounting for sociocultural variations among different peoples.

Environmental Determinism and Evolutionary Theory

Environmental determinism returned to the fore of intellectual consideration of human-environment relations during the late 19th and early 20th centuries following the advancement of Charles Darwin and Alfred Wallace’s theory of evolution. Even prior to evolution, scientists such as Herbert Spencer attempted to apply biological principles to the study of society. Although neither Darwin nor Wallace linked the principle of evolution with the concept of human social and economic development, many others did; they envisioned the potential for what, in their view, would allow for a more rational study of human society and culture. The theory of evolution was embraced as the
bridge linking all organisms (including humans) to the natural laws governing their environment. Friedrich Ratzel, a central figure in the refinement of environmental determinism during the modern period, is perhaps best known as a result of his often-quoted student Ellen Churchill Semple. Semple’s forceful arguments in favor of a strong version of environmental determinism overshadowed the more moderated writings of Ratzel. Anthropogeography, Semple’s brand of environmental determinism, envisioned nature as the supreme agent of influence, molding the physical features of humans and shaping nearly all aspects of humanity’s social development. Environmental factors during this period came to be commonly viewed as the determinative cause of racial differences, cultural practices, moral values, ingenuity, and the ultimate capabilities of any given population.

Criticisms

Despite offering a convenient way to explain difference as a product of geography, environmental determinism quickly fell out of favor following the identification of several fundamental theoretical weaknesses. George Tatham, reflecting on the demise of this theoretical approach, identified three critical problems that ultimately led to its popular demise by the mid 20th century. First, the theory failed to prove that similar environments yielded the same response in human populations, demonstrating that other factors could override the effects of the physical environment on humanity. Second, an increasing awareness of the ability of humans to modify their environment greatly weakened the arguments of environmental determinism given the assumed unidirectional influence of the environment on humans. Last, a consideration of environmental factors alone failed to explain the most fundamental of social characteristics, such as the distribution of the human population.

These failures were further accentuated as early environmental determinists frequently relied on cause-and-effect relationships, established by overly simplistic correlations, to craft their analysis of the environment’s influence on humanity. Additionally, their arguments were further weakened by a reliance on secondary data that limited their conclusions to global scales of generalization. Finally, the deployment of environmental determinism to justify colonial exploits and the subjugation of other societies cast a pall over the entire body of work during this period.

Neo-Environmental Determinism

Renewed interest in the environment has led some to look past the failures of environmental determinism and advance new arguments that attribute greater causal determinacy to the natural environment. Widespread public and academic interest in the work of Jared Diamond, particularly in his Pulitzer Prize–winning book, *Guns, Germs, and Steel: The Fates of Human Societies*, indicates the persistence of environmentally deterministic logic and the failure of human-environment researchers to effectively communicate subsequent theoretical advances. Struggling to account for the distribution of wealth and power around the globe, Diamond forwards his deterministic theory that the conquest of non-Western societies by the West was largely the result of the uneven endowments of the natural environment. The title of the book acknowledges that more immediate factors (guns, germs, and steel) have played an important role; however, Diamond contends that their occurrence is causally determined by the influence of the environment, particularly the growing season of crops. Avoiding the racist conclusions attributed to earlier environmental determinism, neo-environmental determinists reject consideration of the influence of the environment on human characteristics, focusing instead on the effects of the environment on human society. With little substantive difference from earlier determinist perspectives, neo-environmental determinism continues to be plagued by the same theoretical weaknesses as its theoretical antecedent.

*Gabriel Judkins*

*See also* Anthropogeography; Blaut, James; Diamond, Jared; Human Geography, History of; Huntington, Elsworth; Race and Nature; Ratzel, Friedrich; Semple, Ellen Churchill; Social Darwinism; Taylor, Griffith
Discourse is a powerful concept that geographers and others have used to appreciate how nature and the environment—things that are typically thought to be self-evident, presocial, and taken for granted as being outside of, or beyond, politics—are invested with meaning, power, and politics. Thinking of environmental objects (such as a forest, landscape, and nature) as discursive products (i.e., as things made and remade through material and conceptual social labor and connected to deeply situated knowledges and experiences) has greatly opened up geographical understandings of both the “environment” and environmental practice.

The study of the environment as discourse and of environmental discourses (discourses about the environment) has produced many critical insights into the social, political, and historical conditions and processes through which nature and the environment are made meaningful and visible to us as material entities, as well as about the manner in which environmental practices are made meaningful as viable possibilities. These include, for instance, insights into the manner in which concepts of “nature” and “wilderness” have been made and remade through colonial registers of meaning and used as powerful frameworks for dispossession in colonial contexts. The concept has also, in more general terms, enabled critical insights into how the manner in which the environment is made meaningful, incorporated into modes of production, and defined through environmental practice are intimately connected to (and constitutive of) practices of racialization, gendering, sexualization, and class distinction. In moral terms, the study of environmental discourse has helped geographers understand how and why it is—and through what strategies—that environments and the people living in them can come to be conceived in ways that authorize and legitimize their incorporation into regimes of extraction and other modes of production and consumption. For instance, these can include environments as sacrifice zones, places of recreation, or wilderness enclosures.

Further Readings


Conceptual Underpinnings

The intellectual claim that the environment is discursively constructed is grounded in distinct philosophical positions on (1) how the world *is* (ontology) and (2) how we come to know it (epistemology).

1. There is no external or innocent “reality” outside society or social experience, no natural “outside” to culture that can safely be called *nature*, no quality or state that can be thought *natural*. The truth about nature, the environment, and the world is therefore not self-evident or simply “out there,” waiting to be discovered; rather, it is an effect of social practice. In this sense, the environment and nature are “constructed”—encountered, brought into thought, and made meaningful through social practice.
Equally, there is no position outside society and culture (and its modes of production) from which to observe nature (or society for that matter) and make claims about its reality. Knowledges of the environment, about its reality and truths, are always discursive constructions since they cannot escape the socially mediated concepts, meanings, and practices through which we experience the environment and formulate claims about it. Nor can they escape the modes of cultural and economic production that bring knowledge into relation with the environment and already render it visible through technology and economy in particular ways. Furthermore, the idea of discourse incorporates the insight that observers of reality are always positioned in and constituted by social life, so that observations about and knowledge of the environment are contingent on the (gendered, racialized, sexualized, and classed) conditions and concepts according to which they were made. This means that we encounter the environment on particular, partial, and always political terms.

Environmental discourses are therefore the practices that bring into thought spaces of visibility in which environmental objects and our investment in them are constituted, the frameworks through which the environment (and whatever it is juxtaposed to) is rendered legible, and the ideas and concepts that materially alter and determine the state and form of the environment and those whose livelihoods depend on it. Crucially, these are all constructed within particular historical and material conditions and within specific sets of economic, cultural, and technological relations. Environmental discourses are the systems of meaning that bridge how we understand and come to know the world and how the world is as we experience and interact with it. Indeed, as an expression of the relationship between thought and materiality, epistemology and power, the symbolic and real, environmental discourse is a concept that collapses the distinction between the production of meaning and structures of the world.

**Operation of Environmental Discourses**

To clearly understand the concept of discourse, it is helpful to understand how it works—how meaning is made and structured through discourse, how its truths are established, how knowledges are produced, and how social control is exerted. To this end, environmental discourses can be thought of as sets of related statements (including representational practices and performances) about an environmental object or objects that cohere to constitute and produce meanings about it and generate effects around it. To simplify, discourses have (1) objects (things that they constitute) such as the environmental effects of a proposed mine development; (2) effects (things that they accomplish and produce), for instance, to normalize a particular policy direction; and (3) strategies (techniques for establishing the meaning of objects and accomplishing effects).

1. The objects of environmental discourses are the environmental issues, subjects, ideas, and knowledges around which meanings (and more broadly discourses themselves) are made and cohere. For instance, a discourse connected to a proposed and controversial mine development might work to constitute and define objects such as the effects of the mining operations, whose knowledge is appropriate for predicting or determining these effects, and which knowledge-making techniques count as valid. To continue the example, if risk is used as the framework within which effects are constituted and appropriate knowledge determined, the result may be a discourse of risk that coheres around the definition of knowledge, who holds it and who does not, whose claims about the future are valid, how claims about the future should be made, and whether and if control of the future is possible. It is helpful to note here that environmental discourses are dialogical: They exist and change in relation to other, often competing discourses, sets of claims, experiences, and knowledges, as well as other sets of social and political designs. Insofar as they exist in relation to these alternatives, they also incorporate and represent their competing claims and experiences.

2. Environmental discourses also produce material outcomes and effects. They establish the truth about an environmental problem or development, normalize (make to seem natural and self-evident) policy outcomes and particular
knowledges or experiences of the environment, and establish truth about ideas such as “nature” or “the forest.” Environmental discourses, in other words, set the terms within which the environment and environmental problems can be thought; they determine the conditions of possibility for thinking them in the first place; they normalize and legitimate particular kinds of knowledges and experiences about the environment; they make environmental policy outcomes seem natural or legitimate; and they define the environment and nature according to the interests of (usually dominant) groups.

3. Discursive strategies are the material processes and practices through which objects are brought into discourse and through which objects and effects are constituted and defined. Discursive strategies are the devices through which knowledge about discourse objects is developed and through which meanings are conveyed. They are also the material processes and practices through which objects are defined and rendered meaningful. Discursive knowledges about the environment are developed and conveyed, in particular, through processes of normalization and difference making (see below).

Power

Environmental discourses work to realize and make material the interests and experiences (usually) of the powerful. While it is crucial to note that environmental discourses labor and are powerful in and of themselves, they always have authors attached to powerful sets of interests materially situated within uneven geographies of production, consumption, and accumulation. Discourse is itself a mode of production in which power is transformed through constant labor into material expressions on the ground—expressed as and through the subordination and privileging of peoples, the subjugation of environmental knowledges, and the advancement of environmental logics. As a result of the work that environmental discourses do (to define objects and establish outcomes) and of the ways in which they do this work (through practices of normalization and difference making), environmental discourses are powerful. They rely on particular constellations of power, they produce power by laboring to maintain conditions of possibility and constitute their objects, and their effects consolidate and reflect existing power relations.

Normalizing Power and Difference Making

Environmental discourses are contingent. As the practices and frameworks that direct and make possible environmental practices, legitimate modes of production and regimes of extraction, and normalize ways of interacting with the environment, discourses reflect, normalize, and promote particular environmental knowledges, experiences, and epistemologies. Environmental discourses therefore universalize particular experiences and knowledges of the environment that are politically and socially situated. Discourses are powerful insofar as they connect particular (in the sense of partial, limited, local, and embodied) knowledges and experiences to the control of conditions possibility in terms of how the environment can be thought, encountered, mobilized, and used.

Power works through discourse as a result of its ability to hide and obscure the particularity of its constitutive epistemologies, experiences, and politically, historically, and socially situated constellations of meaning. Power also works through a discourse’s ability to make some knowledge claims and systems supersede others. This power to normalize—to present particular social experiences or epistemologies as though they were universal, to hide their particularity from view, and to define alternative experiences as abnormal—obscures the contingency of the discourse. In doing so, it substitutes particularity for universal applicability. Environmental discourses both hide and marginalize environmental knowledges and experiences. The objects and effects of environmental discourses must therefore always be seen to be the results of specific historical practices (including knowledge production) and specific and local environmental knowledges and experiences that arise within particular political, economic, and historical sets of relationships but that masquerade as universal.

Environmental discourses are also difference-making devices, in which the environment and difference are mutually structured. A discourse’s hidden knowledges work as instruments of
normalization around which difference is oppositionally defined. Discourses enact, and produce, registers of difference in relation to environmental objects and use concepts such as “nature” to normalize and naturalize the differential positioning of people in relation to each other in and through the environment. Groups are racialized, sexualized, and defined as deviant against the normal, in relation to how the environment is meaningful through such registers. Discourses produce power as they produce difference (differentiate).

To conceptualize the environment (environmental knowledges, policies, environmental objects themselves, and ideas about the environment such as “nature”) as discursively constituted is to take a critical and political stance toward the operation of power in environmental practice. This stance highlights the importance of deconstruction as a political-environmental project and insists on the possibility of undoing the imperial work that discourses do. Deconstruction aspires to uncover and disable the contingency (and hence the power) of established discourses. As a critical, political project it recognizes that the power of discourse—enabled through the constant work required to maintain and rearticulate conditions of possibility—is susceptible. Despite the considerable authority that environmental discourses exert over our understandings of the world, and despite the manner in which discursive registers are often inextricably entrenched in the meanings through which we navigate the environment, their structures are fragile. If discourses are revealed to be particular, their partiality brought to light, and their disavowals uncovered, their dominance and hold diminish. The operations through which discourses establish their power thus also provide the political opportunities to overturn it. As a critical and political project, deconstruction enables us to see that the environment and our relationship to it might be oriented differently and in more socially just combinations.

Anna Stanley

See also Colonialism; Critical Studies of Nature; Difference, Geographies of; Discourse and Geography; Environmental Imaginaries; Environmental Justice; Nature-Society Theory; Situated Knowledge; Social Construction of Nature

Further Readings


ENVIRONMENTAL ENTITLEMENTS

The Nobel Prize–winning economist Amartya Sen created an entitlement approach to the analysis of hunger and poverty, by which people are entitled to bundles of commodities by virtue of their inherent legal rights. The concept of environmental entitlements is a broadening of Sen’s approach. It is an extension not only in terms of linking economic entitlements to environmental resources, but the environmental entitlements perspective stretches much further, integrating actor-oriented approaches, theories of conflict, power and entitlement distribution, and discursive debates over the meaning of entitlements. In a seminal paper on environmental entitlements, Melissa Leach, Robin Mearns, and Ian Scoones started with a fundamental critique of the then prevailing narratives around community-based natural resource management (CBNRM), an approach that had emerged in the 1990s as a local-level, community-centered, and participatory concept to achieve sustainable development in rural areas of developing countries. The authors initiated their inquiry by asking why CBNRM initiatives had so often failed or fallen short of expectations. The critique centered mainly on the flawed conceptualizations of both village communities and local environments, the categories that are central to the development efforts to support local-level sustainable development. The authors argued that CBNRM frames communities as well as local environments as basically distinct, undifferentiated,
and relatively stable entities. Moreover, a common image underlying people-environment relations in CBNRM is blamed as one of harmony, equilibrium, or balance between livelihoods and natural resources. Interrogating these basic assumptions on communities, the authors proposed an alternative framework that highlighted social difference and dynamics, people’s agency, and the fundamentally contradictory and contested nature of local-level community-environment relations. They spoke to the ways in which gender, caste, wealth, age, origin, and other aspects of social identity divide and crosscut community boundaries. Also, they emphasized that communities are composed of people who actively monitor, interpret, and shape the environments around them. Social change, in this context, has to be framed as manifold processes of interactions between external and internal actors and events, in which path dependencies and contingencies play a significant part.

In a parallel strand of critique, the authors also challenged the conventional framing of local ecologies. On the basis of recent advances in the emerging field of “new ecology,” they argued that a high degree of variability in space and time characterizes local environments and that historical dynamics and disturbance regimes have to be recognized in order to grasp the dynamics of environmental change at a variety of timescales and spatial levels.

Reconceptualizing CBNRM Approaches

Leach and colleagues concluded that local communities are best seen as the temporary and contingent outcomes of dynamic interactions between differentiated social actors, which are frequently manipulated by powerful and by no means shared interests. In parallel, local ecologies have to be framed as dynamic, transforming, nonequilibrium products of both social and ecological histories.

Rejecting the view of communities as static and undifferentiated, seeing the environment dynamically, and disaggregating it into its constituent parts has profound implications for analyzing the link between people and natural resources. In particular, these reconceptualizations raise a very different set of questions from those addressed by conventional narratives around CBNRM: Which social actors see which components of variable and dynamic ecologies as resources at different time periods? How do different social actors gain access to and control over such resources? How does natural resource use by different social actors transform different components of the environment? To systematically address such questions, Leach, Mearns, and Scoones drew on three distinct strands of analysis: (1) an extended version of Sen’s entitlement approach, (2) recent advances in institutional economics, and (3) structuration theory.

From Economic Entitlements to Environmental Entitlements

Sen’s basic concern was to examine how different groups of people gain entitlements from their endowments and thus improve their well-being and their capabilities. Entitlement analysis, from Sen’s perspective, focuses almost exclusively on entitlement mapping—how endowments such as land or labor power are transformed into entitlements—the rights and opportunities to effectively command resources through market channels and legal frameworks. While the shift from supply-driven to demand- and access-driven perspectives on well-being has made Sen’s framework highly powerful for a disaggregated analysis of poverty and hunger, Leach and colleagues considered it to be too restrictive to grasp the real-world relations between differential communities and their resource use. They particularly criticized Sen’s economistic and legalistic ways of mapping entitlements. First, the entitlement approach gives limited attention to endowment mapping, thus ignoring the way which people gain or lose endowments in the long run. Second, the authors criticized the view that entitlements are exclusively distributed through market channels, highlighting the multiple alternative ways of gaining (or being denied) access and control over natural resources. Such alternatives, for example, consist of social networks, social capital, and other informal distribution mechanisms at the community level. Third, they challenged Sen’s legalistic view, considering entitlements as stemming from formal legal property rights only. They pointed to the multiple ways in which entitlements are
distributed through customary laws, community rights, and culture-specific rules and norms.

Given these criticisms, Leach and colleagues defined endowments as the rights and resources that social actors have. More specifically, entitlements refer to legitimate effective command over alternative commodity bundles. Environmental entitlements, then, are defined as alternative sets of utilities derived from environmental goods and services over which social actors have legitimate effective command and that are instrumental in achieving well-being. This definition of environmental entitlements still lacks evidence on how legitimate and effective command over resources is constituted, contested, and struggled over in local communities; how it is gained and lost. From this perspective, environmental entitlements have to be conceptualized as the outcome of conflict and negotiations among social actors, involving power relations and discursive debates on meaning. Leach and colleagues thus argued that an institutional focus is needed for mapping environmental entitlements.

Institutional Perspectives on Environmental Entitlements

Following recent work on institutional economics that conceptualizes institutions as “the rules of the game” in society, Leach and colleagues argued that institutional arrangements and institutional change are central to the environmental entitlements framework. A multiplicity of formal and informal ways that regulate people-environment interactions have to be considered that mediate differential access to and control over environmental endowments and entitlements. As empirical analysis of environmental entitlements points to the fact that rules are constantly made and remade through people’s practices and that many of these practices are embedded in informal institutions, institutional change in local-level communities may be conceptualized as a slow, contingent, path-dependent process. From this perspective, Leach and colleagues defined community-based institutions not as the rules themselves but as regularized patterns of behavior that emerge from underlying structures or sets of rules in use.

This definition points to the problematic relationship between people’s agency, on the one hand, and underlying structural settings, on the other. The authors concluded that actor-oriented approaches in general, and structuration theory in particular, may serve to elucidate how people-environment dialectics in the environmental entitlements framework are played out.

Structuration of People-Environment Interactions

While the environmental entitlements approach, in conjunction with institutional analysis, can serve to conceptualize how social actors access and use environmental goods and services and how this access is embedded in power relations and social conflict, the question remains, In what ways is the environment shaped and transformed through people’s interaction with it? When addressing the dynamic nature of both the social and the environmental constituents of the environmental entitlements framework, Leach and colleagues argued that environments are constantly emerging and transforming as the outcome of dynamic and variable interactions between ecological processes and human resource use. Institutional dynamics shape the ways in which differential social actors perceive, use, and manage components of the environment as resources. From the perspective of structuration theory, the environment provides a setting for social actions, but it is also a product of such action. Environmental entitlements are thus the outcome of dialectic processes between structure and agency in people-environment interactions.

Applications of the Environmental Entitlements Approach

When establishing the environmental entitlements approach, Leach and colleagues had three basic objectives in mind. First, they aimed at improving the conceptual and theoretical base of CBNRM through recent advances in social science and ecology theory. Second, they sought to enhance the social and environmental impacts of CBNRM through a new emphasis on entitlements and capabilities. Finally, they wanted to improve the practice of CBNRM through a particular focus on institutions as mediators of people-environment relations.
In the past decade, most of these objectives were implicitly achieved, in general terms, for community-based approaches to sustainable development, including use of the entitlement framework, institutional analysis, structuration theory, and new ecology approaches. Only rarely, however, was the environmental entitlement framework strictly applied as the basis for analysis. In most cases, the concept was extended or focused in specific ways. One such focus, for example, was on customary marine or inland fisheries systems, with particular emphasis on the empowering and institutional perspectives of managing fisheries. The environmental entitlements concept has also been extended beyond CBNRM and integrated into political ecology approaches, vulnerability analysis, conservation issues, and more recently conflict studies.

In practical terms, the environmental entitlements approach became part of the wider poverty-environment discussion, as, for example, pursued by the U.K. Department for International Development in the context of the Sustainable Livelihoods Framework. Stretching even further, the environmental entitlements approach was explicitly used in official water allocation frameworks. The Ministry for Environment of Victoria/Australia, for example, issued a new Water (Resource Management) Act in 2005 that created the legal foundation for water to be set aside to maintain environmental values of rivers and streams. This law is intended to embed environmental provisions into existing bulk entitlements that will be converted into individual or community-based environmental entitlements.

Criticisms and Prospects

The few preceding examples of how the environmental entitlements framework was applied point out that the concept, with its focus on CBNRM, was probably too narrow in its scope of analysis from the very beginning. In terms of scale, the framework did refer neither to individual nor to larger-scale environmental entitlements, for example, at the district, provincial, or country level. The timescale was also short-term, with little consideration of long-term social or environmental change. Regarding its theoretical and conceptual foundations, the environmental entitlements framework did not really consider issues of social and environmental vulnerability, which were to become so prominent in the Sustainable Livelihoods Framework. Moreover, the epistemologically highly problematic relationship between society and nature, which has been critically examined by recent advances in the analysis of coupled social ecological systems or in social ecology approaches, was treated rather simplistically and superficially in the environmental entitlements approach. More specifically, the extension of the entitlement analysis as presented in the environmental entitlements framework probably did not go far enough too. As Michael Watts has noted, Sen’s conceptualization of entitlements is not suited to grasp socially determined entitlements, nonlegal entitlements, and those dimensions of social well-being that are not based on entitlements at all.

Nevertheless, the environmental entitlement approach, if extended systematically, has a large, widely untapped potential in contexts that stretch far beyond its original objectives and applications. One such potential certainly lies in the vast field of global environmental change and, more specifically, climate change. Recent advances in vulnerability, adaptation, and resilience analysis as crosscutting themes on global environmental change could benefit greatly from environmental entitlement perspectives. The particular focus of the environmental entitlement approach on actors, action, and agency is widely missing in such new systems-oriented approaches to global environmental change. And so is an emphasis on power relations, conflict, negotiations, and institutions in globalized social-nature interactions and dynamics, which the environmental entitlements approach had so powerfully put forward. There is broad scope for an extended environmental entitlements approach to become an essential element in the broader discourse on global environmental change.

Hans-Georg Bohle

See also Common Pool Resources; Community-Based Conservation; Community-Based Natural Resource Management; Environmental Justice; Environmental Rights; Environmental Security; Human Ecology; Hunger
Further Readings


**ENVIRONMENTAL ETHICS**

Environmental ethics represents an emerging area of philosophical inquiry concerning the morality governing human responsibility, valuation, and behavior toward the nonhuman biophysical environment. Traditionally, philosophical approaches to ethics have involved questions of personal, interpersonal, or societal behavior. The search for the moral life inspires questions such as “What is our moral responsibility toward each other?” and “How do we decide which behaviors are right and which are wrong?” This central concern of ethics is to establish an appropriate moral code for living and interacting with others around you. Social ethics are self-reflective; they help determine the guidelines by which we orient our behavior toward others. They inspire, limit, and regulate interpersonal or societal interactions. Environmental ethics is the extension of social ethics beyond the realm of interpersonal human interaction. As such, environmental ethics generate or reflect ideals that limit and inspire human behavior regarding the physical environment and ecological systems.

In the United States, the practice and systematic inquiry now termed environmental ethics began to emerge during the growing environmental awareness of the 1960s and 1970s. Seminal writings by Rachel Carson, Garrett Hardin, and Lynn White resonated with the widespread public realization that large-scale, rapid, and perhaps irreversible changes to physical environments and ecological systems were causally linked to people. Human behavior, values, and attitudes became the problem—they were the causes of the environmental crisis. One outgrowth of this “Green Revolution” was a new form of environmental activism that led to the first Earth Day on April 22, 1970, and a body of federal legislation highlighted by the National Environmental Policy Act (1970), Clean Air Act (1970), Clean Water Act (1972), and Endangered Species Act (1973). Another outcome was a loss of faith in the dictum of modernistic progress and a desire to challenge the set of moral rules condoning patterns of human behavior contributing to environmental and ecological crises. To do so would require diagnosing why traditional ethics allow and even justify widespread degradation and despoliation and systematically eliminating these elements from a pro-environmental conception of ethics.

In response, writers and philosophers began turning their attention to reconstructing normative modes of morality and behavior that would not only counter but also reverse human-induced degradation of environmental and ecological systems. The goal for many of these ethicists was to establish a moral code that could impose limitations on human freedom of action with respect to the Earth. But far from viewing environmental ethics as a purely philosophical enterprise, practitioners always looked to apply these ideals to help solve environmental crises. J. Baird Callicott and others explicitly linked the products of their ethical theories to the goal of influencing human behavior at the structural scale as well as the personal scale. The immediacy of environmental problems led many activists to employ these ethical formulations in justifying emerging environmental laws and regulations, giving their ideas tacit institutional authority.

Modernism, scientific progress, and human exceptionalism came to be understood by environmental ethicists as separating humans from nature. This philosophical dualism, grounded in Western traditions dating back to Aristotle and
reasserted by René Descartes, lies at the root of environmentally destructive behavior because natural environmental systems are largely defined in opposition to people and society. As a counter-weight to Cartesian dualism, much of the work within environmental ethics seeks to emphasize a more holistic, ecological approach.

Whereas social ethics tends to stress atomistic thinking by giving primacy to individual rights and claims against other individuals, environmental ethicists seek to reorient moral claims around collective rights and responsibilities toward the Earth. Monistic arguments take holism a step further by asserting that humans are an animal species that has developed and evolved within the ecological context of a historical Earth. Such fundamental questions arising over whether humans and the imprint of cultural practices and religious beliefs on the physical landscape are natural, supernatural, or otherwise distinct from nature led to a series of ongoing debates about how to categorize and assign value to landscapes and ecosystems that are, to some extent, artifacts of human action. Terms such as nature, second nature, false nature, and wilderness have become complex concepts embodying these debates over the separation of people from the nonhuman “other.”

At the core of many debates that have been generated within environmental ethics is the disharmony present in attempts to balance the interests of people, nonhuman life forms, and the underlying physical and chemical processes creating and supporting the habitats of both. Who or what has rights that need to be respected by humans and human behavior? Much of the fragmentation within the field of environmental ethics can be traced to the different ways of answering this question and justifying it philosophically.

Anthropocentric arguments align more generally with traditional philosophical positions claiming that either only humans have a moral standing or, if it is to exist outside of people, it is contingent on rights being granted by humans. Accordingly, it is not immoral to cut down a tree, kill a snake, or drain wetlands if it is deemed to be in the best interests of humans to do so. Anthropocentric ethics are not necessarily anti-environmental since protecting forests, species, and biotic communities may also be deemed to be moral if it is in our best interests to do so. In both scenarios, people determine worth and are the source of moral value.

The concept of intrinsic value breaks down these traditional boundaries limiting rights to individual humans. It bridges the liberal extension of rights to oppressed or powerless people to other living things or even biotic communities. If an object, animal, or place has a moral standing beyond that granted by people, ideals of ethical behavior include an obligation of respect to these phenomena. Peter Singer argues that a refusal to extend moral rights to animals is a form of speciesism, similar to forms of racism or sexism that deny rights based on race or sex. An unrelated but similar vein of thought, ecofeminism, equates the subjugation of nature with the patterns and justifications made for the subjugation of women by men. Other writers such as Tom Regan contend that other species have some value in and of themselves that is independent of that which we are willing to grant. Such biocentrist have sought to reorient ethical mores on the conviction that all living things have intrinsic value and cannot be excluded from moral consideration.

One of the most influential authors on environmental ethics, Aldo Leopold, predated the formal development of the field itself. Leopold’s conception of the “land ethic” has been seen as the blueprint by which intrinsic moral value can be extended to the synergisms of ecology. Ecocentrism extends moral consideration past either anthropocentrism or biocentrism. Not only living individuals contain intrinsic value; the integrity, stability, and beauty of the biotic community have value as well. The complex webs of organisms, chemical fluxes, and physical processes generating and sustaining ecosystems is the locus of intrinsic value, not an individual human. Therefore, behavior or public policies that degrade or despoil biotic communities for the benefit of individual people, social groups, or even the human species are morally unjust.

Michael A. Urban

See also Commons, Tragedy of the; Deep Ecology Movements; Ecofeminism; Ecological Justice; Environmental Discourse; Environmental Entitlements; Environmental Justice; Environmental Rights; Ethics, Geography and; Nature-Society Theory; Social Construction of Nature
Further Readings


ENVIRONMENTAL FOOTPRINT

See Ecological Footprint

ENVIRONMENTAL HISTORY

There are two immutable facts that underlie human-environment interactions. First, humans are part of the biosphere, and we rely on it to meet our essential needs. Second, through human history, but increasingly over the past 300 yrs. (years), we have become the dominant agent of environmental change. Understanding the relationship between humans and their environment has become increasingly important in recent decades due to human-accelerated environmental change, most pronounced in global climate change, biological extinctions, and soil and water degradation.

Environmental history is the field of study concerned with the systematic understanding of long-term transformations of landscapes, ecosystems, and natural resources through the interaction of humans with the environment, and the role of feedbacks of environmental change on human societies (Figure 1). Environmental history aims to provide the context for understanding our present environment, indicate what past environments were like, and identify lessons we can learn for building a sustainable future. The key elements of environmental history revolve around (a) analyzing human-modified landscapes through the linkage of socioeconomic and biophysical data, explaining how historical processes such as human settlement patterns and disturbance regimes help shape contemporary landscapes and ecosystems; (b) addressing the temporal and spatial dimensions of human-landscape transformation; and (c) understanding human-modified landscapes as cultural legacies.

Environmental history has discarded former dualistic approaches to human-environment relations, instead conceptualizing that social and natural systems are inextricably intertwined. Equally important, environmental history challenges the notion of pristine landscapes or “natural systems” with the belief that it is difficult to find a landscape or ecosystem anywhere on the planet that has not somehow been changed by human activities. As human impacts on the global environment continue to accelerate, environmental sustainability has become an increasingly important part of the global social conscience and international policy agenda. In this context, environmental history is an essential component of environmental planning and management.

The following sections outline the emergence and development of environmental history as a discipline, the methods used by environmental historians, and some common issues in the study and application of environmental history. The final section considers the application of environmental history to contemporary human-environment interactions.

Development of Environmental History

Over the past 200 yrs., there has been a gradual recognition of the power of humans to transform whole landscapes and ecosystems through land use. Explorers, travelers, and natural historians
such as Pierre Poivre and Alexander von Humboldt at the turn of the 1800s saw the destructive impacts of colonialism and the consequences for tropical landscapes. In 1854, the American Henry David Thoreau published *Walden*, in which he reflects on the problem of land transformation during the Industrial Revolution, adopting a political and philosophical perspective. In 1864, George Perkins Marsh published *Man and Nature*, one of the first books to analyze the impacts of human activities on the environment, comparing environmental changes in the Roman Empire with those in the United States at that time. In the 1920s, the cultural geographer Carl Sauer theorized also that different cultures developed distinctive cultural landscapes, rejecting the idea that the environment determined different cultural characteristics.

Two important origins of environmental history can be distinguished in the 19th and early
20th centuries: one from ecology, the other from geography. Ecology emerged from a historicized field of biology and geology in the 19th century, with the first ecology textbook written by Eugen Warming in 1895. The maturing of ecology in the early to mid 1900s became a source of inspiration and instruction to environmental historians by bringing an increasingly holistic framework to often scattered and narrative data and evidence. This was especially evident from the birth of Arthur Tansley’s concept of an ecosystem in 1935 and the emergence of Frederic Clements’s theory of succession to characterize long-term vegetation dynamics driven by the interaction of ecological and human management factors. The consolidation of plant ecology highlighted the importance of vegetation as an indicator of biophysical and management processes and its potential use in historical studies of the environment. In 1949, the American ecologist Aldo Leopold called for the development of an ecological interpretation of history.

Geography provided an increasing awareness of the spatial dimension of social and ecological processes and their interaction, with historical geographers documenting these relationships in many parts of the world. By the mid 20th century, however, geography became focused more on human activities, and the environment formed the backdrop to these activities rather than a central interest. In the 1950s, the German geographer Carl Troll pleaded for a more holistic conception and approach to the study of human-environment relationships. Troll understood the advantage of linking geography and the biological sciences and introduced the new discipline of landscape ecology, in which humans are considered an integral component of landscapes and ecosystems. The first half of the 20th century could be called the early period of the development of environmental history, with no formal disciplinary recognition.

Environmental history as a specific discipline emerged in the United States in the 1960s and 1970s, parallel to the rise of the environmental movement. People around the globe became aware of and concerned about the negative impacts of human-accelerated environmental change. Factors such as population growth, industrial and technological change, and the globalization of trade and consumption were exhausting natural resources at an unsustainable rate. Societies recognized that not only was there a moral responsibility to minimize or mitigate the negative environmental consequences of our actions on the biosphere but humans were ultimately dependent on the natural world for their survival. By learning from history, the hope is that sustainable land management outcomes will arise.

The first course specifically concerned with environmental history was introduced at the University of California, Berkeley, in 1970. It centered on a long-standing interest in documenting the American experience with wilderness, with the environmental change as evidence of people’s attitudes and actions. In 1976, the American Society for Environmental History was founded, while the first Institute for Environmental History in Europe was created in the 1980s, based at the University of St. Andrews in Scotland. In the mid 1980s, the environmental historian Donald Worster emphasized the need for an ecological perspective in historical study, refining this into an agroecological perspective. In fact, the term ecological history is often used interchangeably with environmental history. Worster stressed the need to understand the organization and functions of nature, how modes of agricultural production have altered human relationships with their environment, and the role played by human attitudes, values, and perceptions in transforming the Earth. The European Society for Environmental History was not organized until 1999, but it built on a strong tradition of historical geography research in Europe. The recently created International Consortium of Environmental History Organizations brings together American, European, and Australasian societies of environmental history. Interest in environmental history has followed different trajectories in different regions, with strong development in places such as the United States, Europe, Australia, and Africa but less so in Russia, the Middle East, and Japan.

Environmental history is increasingly becoming a meeting ground of historians, geographers, ecologists, biologists, anthropologists, and conservationists. Indeed, many scholarly contributions that are pertinent to environmental history are written by scholars who typically would not
identify themselves as historians. Topics range from human relations with wilderness through the development of agricultural lands, the use of different natural resources such as water or timber, and the impacts of colonial expansion and settlement of “new lands” to providing a basis for the holistic management of landscapes. There is a strong bias toward terrestrial environmental history, with little current work on marine areas.

In recent years, as a result of growing concerns about the accelerating rate of global environmental change, environmental history has transcended the academic arena and become a popular topic, with books such as Jared Diamond’s *Collapse* and *Guns, Germs and Steel*, Alfred Crosby’s *Ecological Imperialism*, and *1491*, Charles Mann’s synthesis of the reality of the extent and characteristics of human impacts on the landscapes of the Americas before European settlement.

### Environmental History Methods

While environmental history can be constructed in myriad different ways, its distinguishing feature is the integration and synthesis of qualitative and quantitative data from a variety of sources to build a better understanding of the drivers and impacts of past human activities on their environments and the way in which those environments have in turn shaped societies. It is this transdisciplinary nature that sets environmental history apart from mainstream history, geography, or ecology.

Environmental history advocates the understanding of the human-environment relationship in both time and space to gain a holistic perspective of environmental change and its cumulative effects. Temporally, environmental history forms part of a continuum between paleontological studies and the present day and can span timescales from decades to millennia. As both human cultural traits and environmental systems are dynamic, neither is environmental change unidirectional, nor are the outcomes the same for different regions of the world. In addition, the effects of changes arising from human-environment interactions may take many years to become apparent. It is important, therefore, to identify key ecological stages and historical periods during the time period under consideration, together with their underlying processes and drivers.

Spatially, studies in environmental history can be conducted at scales ranging from local to national or regional to global and may focus on specific communities, localities, or occupations or specific resources such as water, fisheries, or agricultural land. However, even if a particular scale such as the region is the focus of the environmental history, consideration of external factors at broader scales is crucial as each factor is influenced by those around it as part of the hierarchy of influence and decision making. For example, individuals create new agricultural fields, but the decision to develop a region is often the responsibility of regional or national governments. Similarly, the type of agriculture may be influenced by international trade opportunities.

Environmental histories are in general narratives or stories about a locality or region, or even the whole Earth, through time. One of the strengths of environmental history lies in its narrative synthesis, which helps deepen our understanding of human-environment interactions, but it is important that the narrative be grounded in an ecologically sound base. Environmental history must work with and integrate heterogeneous data sources, ranging from qualitative accounts to quantitative scientific and statistical databases from different disciplines. It is common that the data that interest environmental historians are highly variable in quality and reliability as well as in their spatial and temporal coverage. This necessitates comparison of disparate sources of data to see if commonalities are present before relying on these data.

A variety of techniques are combined to establish ecological and environmental histories, depending on the focus of the particular study. Some techniques are spatially accurate and precise, documenting environmental change using data such as climate records, aerial photographs and maps, ecological studies, geology and sediment records, or economic development using statistical records of farm production or population growth. Others are more subjective interpretations or “stories” garnered from texts, memories, myths, and personal experience. These may include diaries, oral histories, government reports, or newspapers.

One major limitation for environmental history is the shortage and variable quality of long-term
quantitative data sets. However, especially from the 1990s, the study methods have become more quantitative due to the increasing availability of spatially explicit and quantitative data from remote sensing and geographic information systems, which has allowed the application of sophisticated modeling techniques, linked to those used for studying land use and land cover change. Maps and spatial representations of change are becoming increasingly important in environmental history. One of the latest advances in measuring and analyzing the impact of humans on the environment is the “ecological footprint” concept and its spatially explicit derivations such as the spatial footprint or the wilderness index.

There are several other issues to be aware of when developing or using environmental histories. Most important, the complexity and dynamic nature of both environmental and cultural systems over time mean that even a very detailed environmental history can tell only part of the tale. Second, which part is told depends on both available data and the expertise and background of the researchers, and often their political and ideological beliefs, since environmental historians are embedded in their own cultural and knowledge system. It is, therefore, possible that different and sometimes even contradictory conclusions are reached about the same event. In addition, the ability of a well-written story to sway opinion and become accepted as fact presents a danger that continuing research on an issue, or the development of alternative ideas, may be hindered. This does not mean that environmental histories are unreliable but that, as with any research, as new information becomes available the conclusions should be updated. These issues aside, environmental history has deepened our understanding of how past societies have interacted with and changed their environments. Increasingly, this knowledge is seen as underpinning environmental planning and management in a number of arenas.

Examples and Applications of Environmental History

As mentioned earlier, recognition of the unsustainable cumulative impacts of human activities on the biosphere led to the development of environmental movement and, from this, environmental history.

While environmental history, by its very nature, focuses on the past and allows us to know the ending of the particular story, the hope of most environmental historians is that it can also alert societies to a range of possible outcomes from present-day activities. “Big picture” environmental histories that document global or continental change have increased our understanding of the connectivity between human and environmental systems in both space and time. For example, the introduction of often highly invasive exotic plants and animal species associated with colonialism during the 18th and 19th centuries has left a permanent environmental legacy in many countries and continues to be a major contemporary threat with the increasing globalization of travel and trade. There are also common traits in human-environment interactions in the settlement of new lands, from the expansion of Europeans into the Americas, Africa, and Australia to the spread of Polynesians across the Pacific. Resources in new lands were often seen as limitless, and there was little understanding of the environmental impacts of settlement on native ecosystems. Landscapes were altered dramatically to conform with the different cultural and production systems of the colonizers, leading to rapid environmental change, but the colonizers in turn had to learn to adapt to different environmental conditions. Understanding the commonalities of settling new lands may guide more environmentally sensitive settlements in regions with current development frontiers, such as the Amazon basin.

An area of growing interest to environmental historians is the quantitative analysis of contemporary environmental change focusing at the landscape scale, where aerial photography and satellite imagery can be used, for example, to quantify landscape dynamics in response to both deforestation and reforestation. Studies of landscape changes are increasingly being used to inform regional conservation planning and global change policy.

Environmental history is increasingly being used as a landscape management and planning tool. This is based on the tenet that it is necessary to understand the legacies of past cultures on contemporary landscapes in order to manage and plan for a sustainable future. Even today’s wilderness areas are a legacy of human decisions and
planning whereby they were protected from development, intentionally because of their ecological and ascetic values or unintentionally because of biophysical constraints. Some wild areas are not “wilderness” at all, as in the case of the New England forests, which have regrown because of land abandonment in the mid 19th century, or even large areas of the Amazon rain forests that were cleared and densely settled at the time of European colonization. Because of the slow recovery of many vegetation communities, land use legacies can be extremely long lasting. Although often not explicitly stated, a common goal of environmental historians is to improve the relationship between people and the environment so that ecosystem functioning is maintained or restored as far as possible. To this end, environmental histories are a required stage for ecological restoration projects, identifying the components and processes required to rebuild ecosystem functions.

It is important not to overgeneralize about the relationship between humans and the environment. Just as culture varies among societies, the nature, extent, and intensity of environmental change also vary among societies over time. There was a tendency to see pre-European landscapes in North and South America and Australia as pristine wildernesses, but increasingly there is recognition of the modifications and transformations carried out by indigenous societies and how their cultural systems adapted to and interacted with their environment. Some traditional societies, such as the Australian Aborigines, have managed to live within the capacity of their environment for centuries, while others radically changed their environment within a short time span, such as the Maori in New Zealand. Yet others were unable to survive environmental change, with civilizations such as the Mayans in Central America disappearing at least in part due to drought and exhaustion of natural resources.

Conclusion

Environmental history has steadily evolved as a discipline over the past 200 yrs. Today, it is highly relevant for understanding legacies of past environmental change and guiding conservationists, scientists, policymakers, and land managers in understanding how the past shapes the present and the present may shape the future. Environmental history is, therefore, an important ingredient for land use and environmental planning and policy in the era of climate change and human-accelerated environmental change.

Leonie Seabrook, Andres Etter, and Clive McAlpine

See also Diamond, Jared; Ecological Footprint; Environmental Determinism; Environmental Discourse; Environmental Ethics; Environmental Imaginaries; Environment and Development; Global Environmental Change; Historical Geography; Human Dimensions of Global Environmental Change; Human Geography, History of; Landscape Ecology; Land Use and Cover Change (LUCC); Land Use History; Prairie Restoration

Further Readings


The term environmental imaginaries is used to explain how the natural environment shapes the attitudes, discourses, and practices of the people who dwell there. Derived from the broader concept of a spatial imaginary by Richard Peet and Michael Watts, an environmental imaginary describes how people draw on the familiar elements of their surroundings to understand natural and social processes and to inspire creative ways of shaping those
same processes. It has a strong regional component because of how the natural environment differs from region to region, and it is also closely tied to political ecology with its focus on economic and political institutions.

The environmental imaginary concept has a lineage that can be traced back to spatial and social imaginaries. Developed within sociology, a social imaginary describes how a particular society conceives of itself as well as how it explains both the world around it and its own characteristics. It can be thought of as a bridge between doctrines or discourses and embodied practices. The term imaginary can be interpreted in three ways. First, there is the fact that social imaginaries are implicit rather than explicit, so they exist within people's heads. Second, individuals use them as jumping-off points in their own ways of coping with the world and forming their own identities: Imaginaries spark creativity. Finally, imaginaries are used to apply processes that are already understood to the unknown future. An “imaginary” therefore does not only explain what has already happened but also provides a framework to conceive of the future.

In addition to factors such as political structures, existing institutions, and social makeup, all social imaginaries are shaped by the places in which they are developed and in turn shape those places. Spatial imaginaries describe how a society understands the spatial processes that it creates and encounters, as well as how spaces and places are constructed and made meaningful. Globalization is one of the best-known spatial imaginaries, with lively debate ensuing about whether it is an external force being imposed on helpless local places or constituted by millions of individual or local places and practices. At other scales, businesses use their spatial imaginaries to determine the best places to site a production facility or understand why their goods and services are or are not selling in a particular market; governments at all levels promote economic and social development with regard to their understandings of how processes occur across space; and individuals arrange their lives around spatial patterns of work, home, and recreation. These spatial imaginaries are all fluid and flexible as the actions taken lead to new or changed understandings of how spatial processes work.

The environmental imaginary, as developed by Peet and Watts, is even more explicit about the role of place, arguing that a society's concept of relationships between humans and the environment—and therefore how people act on their environment—varies based on the characteristics of that environment. Interactions between the human and nonhuman shape ever-changing environmental imaginaries. Because of the influence of political ecology, environmental imaginaries are often conceptualized as being about access to and the meaning of natural resources based on political structures as well as biophysical characteristics. Differentiations based on race, class, and gender also enter into the potential creation of multiple environmental imaginaries within the same space.

Peet and Watts use the concept of an environmental imaginary as a pivot point for change and transformation. They draw heavily on the creativity implied in the concept, arguing that liberation ecologies and social movements have to rely on people's conceptions of their own environment in order to be successful. In particular, because environmental imaginaries are constructed based on the physical environment of a particular place or region, they are not meant to be applied elsewhere, which often happens within the context of development programs. Disaster can result when, for example, actors with an environmental imaginary developed in the humid First World tell actors in an arid zone of the Third World that their farming practices are degrading the soil without taking into account existing local practices based on centuries of inhabiting that place.

Somewhat ironically, considering its origins in a book about the developing world, the environmental imaginary concept has been almost exclusively applied to the global North. This includes the transfer of acacia trees in and out of Australia, environmental conflicts in Central Appalachia in the United States, understandings of sustainable development in Australia, invasive species in Toronto, and the history of colonization in Canterbury, New Zealand. Much of this increased attention in recent years has been from the strand of geographic literature seeking to reconcile the human and nonhuman (or “nature”) as well as urban political ecology, both of which are largely Western-centric. Still, significant potential exists to expand the concept within situations outside the developed world.
While there might be reluctance to engage with environmental imaginaries as a theoretical framework because of their narrow focus on the region as a key scale, the increasing interest in bioregionalism from a management and policy point of view (such as watershed management, regionally based green building criteria, and local food movements) makes them more pertinent than ever. Environmental imaginaries offer a way of thinking about the relationship between the human and nonhuman environments that includes both social and physical processes, as well as a means of explaining the past and laying out possible progressive futures.

Julie Cidell

See also Bioregionalism; Deep Ecology Movements; Ecofeminism; Ecological Imaginaries; Environmental Determinism; Environmental Discourse; Feminist Environmental Geographies; Feminist Environmentalism; Geographical Imagination; International Environmental Movements; Nature-Society Theory; Peet, Richard; Political Ecology; Political Economy of Resources; Social Construction of Nature; Watts, Michael

Further Readings


Environmental impact assessment (EIA) is one of the most widely practiced environmental management tools in the world. The International Association for Impact Assessment, the world’s leading authority on impact assessment, defines EIA as the process of identifying, predicting, evaluating, and mitigating the biophysical, social, and other relevant effects of development proposals on the environment prior to taking major decisions and making commitments. As a formal, systematic, and often regulation-based process, EIA is not to be confused with related environmental studies, such as environmental site assessments, whose purpose is to determine the nature and extent of environmental contaminants at a specific site and to identify remediation plans. These latter studies are much more common and shorter, though in some cases they lead to the preparation of an EIA.

The term environmental impact assessment is often used interchangeably with environmental assessment (EA) or impact assessment (IA). Regardless of the terminology used, EIA, EA, or IA refers to an organized and interdisciplinary process designed to gather information used to identify, understand, and manage the potential effects of proposed developments and actions, both public and private, on the biophysical environment (e.g., air, water, land, plants, and animals) as well as on the human environment (e.g., culture, health, community sustainability, employment, financial benefits).

EIA was first introduced in the United States in the late 1960s and is now practiced in more than 100 countries. The environmental movement of the 1960s played a key role in the birth of EIA. The post–World War II era in North America was marked by an unprecedented period of economic growth and also a period of rapid environmental change and increased public awareness of the widespread negative effects of development on forests, wetlands, and water and air pollution. Inspired by pioneering works such as Rachel Carson’s Silent Spring, which described
the adverse effects of pesticide use on the environment, the environmental movement led to increasing pressures on central governments to take actions to address the negative environmental changes that were occurring. The response was the passage of several laws, each addressing a particular environmental concern, such as the Clean Air Act, designed to establish pollution limits and to regulate industrial emissions.

It was not until 1969, when the U.S. National Environmental Policy Act (NEPA) was introduced, that a framework was provided for the simultaneous consideration of the full range of environmental considerations associated with a proposed development activity. The term *environmental impact assessment* was coined to refer to this framework and the process that it generated. NEPA, which became law in 1970, was not designed to replace other laws and environmental programs but to complement them by providing the means to integrate environmental and public concerns about often controversial development proposals and to organize and present these concerns in a way that was meaningful to decision making. For the first time, those proposing to undertake certain development projects had to demonstrate that the projects would not adversely affect the environment and to document this in the form of an environmental impact statement (EIS)—an often lengthy and highly technical document that describes the proposed development or undertaking and reports on its potential environmental and socioeconomic effects and prescribed management measures. During the first decade of NEPA, approximately 1,000 EISs were prepared annually in the United States. This number has greatly increased since then.

### Purpose and Objectives

The primary purpose of EIA is to facilitate the consideration of the environment in planning and development decision making and, ultimately, to make it possible to arrive at decisions and subsequent actions that are more environmentally sound. The specific objectives of EIA can be separated into output objectives and outcome objectives. The output objectives of EIA are the immediate, short-term objectives or returns of applying EIA and include

- improvements to the environmental design of the proposed developments;
- explicit integration of environmental factors in decisions about development actions;
- an opportunity to anticipate, avoid, minimize, or offset potentially adverse environmental impacts before a development becomes a reality; and
- provisions for public debate about a proposed development action.

The outcome objectives of EIA are longer term and the product of consistent and rigorous EIA application. They include

- protecting the productivity and capacity of human and natural systems and ecological functions,
- facilitating learning and environmental education, and
- promoting sustainable development.

### EIA Process

EIA can be thought of as a framework or process to systematically examine the potential environmental implications of development actions prior to their approval. The specific components and application of EIA are dictated by the specific issues it attempts to address and by its regulatory or policy requirements. However, the EIA process stemming from NEPA and subsequently diffused throughout the world requires a series of systematic steps:

**Project description.** Prior to an EIA, the proponent prepares a detailed description of the proposed action and the project environment and a statement of the need for and purpose of the undertaking and its alternatives. This is often called the “application for development.”

**Screening.** The project information is screened against a list of projects that require assessment, regulatory or legal triggers for assessment, criteria and thresholds, or some combination of these to determine whether an EIA is necessary. The screening process ensures that unnecessary assessments are not carried out and that developments warranting an EIA are not overlooked.
Scoping. The proponent or government agency responsible for carrying out the EIA determines, in consultation with the affected public and various interest groups, the important issues and parameters to be considered in the assessment and establishes the spatial and temporal boundaries for the EIA. Attention is focused on identifying what studies need to be done and then establishing the current environmental baseline and socioeconomic conditions of the project region. Potential interactions between the project and a select few priority environmental components, often referred to as valued ecosystem components (VECs), are identified.

Impact prediction. The likely changes or effects caused by the project on the biophysical and human environment are predicted, and the potential impacts emerging from those effects are identified. Consideration is also given to the significance of effects and impacts and to the identification of those effects and impacts that require responsive management.

Impact management. The purpose of impact management is to find ways to reduce, eliminate, or avoid those potentially adverse effects and impacts that are likely to be caused by the development and to identify opportunities to create or enhance potentially positive ones. Impact management may include activities such as developing impact management plans, environmental protection plans, or compensation packages for those effects and impacts that cannot be avoided.

Draft EIS preparation. A draft EIS is prepared describing the project, the current baseline conditions, key issues and likely project impacts, and the proposed impact management strategies and programs for longer-term environmental monitoring.

Public and technical review. Under most environmental assessment systems there is an opportunity for a formal public and technical review of the EIS and related documentation prior to the preparation of the final EIS and a decision being made.

Final EIS preparation and decision. The results of the review are considered by the proponent and integrated into the final EIS. The EIS is then submitted to the responsible authority, which is usually a government environmental agency or department, for a decision as to whether the proposed action should proceed and under what conditions. Under some environmental assessment systems, the responsible authority may conduct a second, independent public or technical review of the EIS prior to making a decision. If the EIS is rejected, it is sent back to the proponent for revision to address the identified deficiencies.

Implementation and follow-up. Once a decision is made for approval, the project is implemented along with any proposed impact management and mitigation measures. Monitoring programs are put in place and continue throughout the life cycle of the development project to provide feedback on the effectiveness of impact management and to identify any new or emerging environmental and public concerns.

EIA Beyond the Project

Although NEPA was never intended to be restricted to individual development projects, such as bridges, dams, and mine sites, EIA emerged throughout the 1970s and 1980s with a very strong project emphasis. It is only since the early 1990s that attention has turned to applying EIA above the project level, to address the potential environmental implications of policies, plans, and programs. This higher-order form of assessment, known as strategic environmental assessment (SEA), is gaining international attention as a means to address environmental impacts earlier in the decision-making process. The rationale for SEA is that development projects are often the result of policies, plans, and programs, or lack thereof, playing out on the ground. Thus, to capture the sources of environmental impacts, there is a need to apply environmental assessment before development projects are proposed. This, in turn, should lead to better project proposals and decisions about development.

Enduring Concerns

EIA has evolved considerably since NEPA. It is now applied on a global scale, occurs earlier in the decision-making process, and is focused more on broader sustainability concerns. However,
several significant issues concerning the practice of EIA exist. These include the ability of EIA to predict the future accurately, the role of the public in EIA, follow-up and monitoring of projects after they receive approval, and the cost-effectiveness of the EIA process.

Central to EIA is the prediction of environmental impacts. This requires some ability to foresee the future—an almost impossible task in a highly complex and changing environment. As a result, impact predictions are frequently inaccurate, resulting in public mistrust of the EIA process and of the science involved. A major challenge is that the impact predictions in an EIS rarely include any discussion of uncertainty or error about those predictions or indications of the quality of the data used. This is understandable for a development proponent in that the public and decision makers are often reluctant to accept uncertainty. At the same time, precise predictions about the future are often suspect, prone to criticism, and more likely to be wrong.

Public involvement has been a long-standing challenge to EIA. The standard approach has been to prepare an EIS and release it to the public for review and comment rather than to involve the public early in decision making, at a stage when the need for and the purpose of the project are being discussed. Early public involvement reduces potential conflict at later stages of the process and brings to EIA a host of public issues and concerns about development that may otherwise be missed.

EIA is often perceived as a means to an end, rather than a long-term commitment to environmental management. Once a development is approved, an EIS is often shelved, and little use is made of the results for postimplementation monitoring and follow-up. Without an ongoing commitment to monitoring, little is known about the actual impacts or effectiveness of the measures implemented to manage them. When monitoring occurs, it is often limited to those components of the environment for which ongoing data are required to renew project or operating permits and licenses.

Many development proponents argue that EIA is not worth the financial investment. Often, EIA and the preparation of an EIS are seen as hurdles that must be overcome to secure a development permit. The result of this view is an EIA process that is separated from project management and ongoing improvements in environmental performance. This is not consistent with the intended purpose of EIA. An EIA presents a significant opportunity for a proponent to save money by minimizing public conflict and opposition to the project, thereby reducing costly project delays, and facilitates overall improved project environmental performance.

Bram F. Noble

See also Environmental Impact Statement; Environmental Impacts of Agriculture; Environmental Impacts of Cities; Environmental Impacts of Manufacturing; Environmental Impacts of Oil Fields; Environmental Impacts of Pipelines; Environmental Impacts of Roads; Environmental Impacts of Tourism; Environmental Management; Open-Pit Mining; Public Policy, Geography of; Sustainable Development; Sustainable Development Alternatives

Further Readings


ENVIRONMENTAL IMPACTS OF AGRICULTURE

Global environmental systems have been profoundly affected by agriculture throughout the Holocene epoch, but these impacts have been especially pronounced since the Green Revolution began around 1945. The environmental effects of food production include alterations of the Earth’s hydrologic cycle, increasing levels of...
atmospheric greenhouse gases, decreased biodiversity, accelerated rates of soil erosion, and the rapid spread of eutrophication in freshwater and marine ecosystems.

The Agricultural Revolution

Crop plants were first domesticated around 10,000 yrs. (years) ago, when centers of crop domestication emerged independently in the Eastern Mediterranean, Asia, Africa, and the Americas. The three staple grains that today account for more than 60% of all caloric intake—wheat, rice, and corn—were all domesticated by ca. 7,000 yrs. BP (before present). However, the domestication events themselves were inconsequential in terms of immediate environmental change. The pervasive and lasting imprints of the agricultural revolution on the global environmental commons did not begin to develop until ca. 4,000 to 5,000 yrs. BP, when agriculture was rapidly emerging across the globe as the primary means of food procurement.

Forest Clearance: Ecosystem and Atmospheric Impacts

Between 5,000 and 3,000 yrs. BP, large cities (10,000 to 100,000-plus inhabitants) were established on several continents, and this growing urban trend was fueled by the conversion of the wooded hinterlands into agricultural fields. Fire was used extensively as a means of clearing both dead and living biomass in preparation for the planting of agricultural plots; stone and metal tools (i.e., axes, metal plows) were also adopted for agricultural purposes, particularly in the Near East, Europe, and Asia. The oft-cited clearing of oak woodlands throughout the Mediterranean region during classical times led to increased soil erosion, which was made worse by the proliferation of domesticated grazing animals such as goats and sheep. The process of stocking domesticated animals on areas that were formerly forested has been repeated on every agricultural continent, often with the same end result: retardation of vegetative succession and the maintenance of the treeless landscape.

The decreases in natural vegetation cover lost to agricultural conversion have resulted in extensive native habitat loss and forest fragmentation and have contributed significantly to modern extinction rates that are several orders of magnitude higher than those experienced during pre-agricultural times. Since agricultural landscape patchworks have been carved out of formerly contiguous tracts of native vegetation, large migratory and wide-ranging animal species have often been relegated to habitat patches too small to satisfy their dietary and breeding requirements. Farmers have also historically viewed native animal species, both herbivores and carnivores, as competitors or threats to successful food production. These “agricultural pests” have long been targets of local extermination efforts and include animals in the United States such as prairie dogs, gophers, wolves, and foxes. In summary, the net effect of agriculture on global ecosystems has been one of biological simplification or decreasing biodiversity.

The rise of agriculture has also precipitated a marked increase in anthropogenic greenhouse gas emissions. The conversion of native ecosystems—forests, shrublands, and grasslands—to agricultural fields has resulted in increased carbon dioxide (CO₂) fluxes from the biosphere to the atmosphere. Some scientists argue that this process of agricultural conversion led to climatically effective anthropogenic greenhouse gas emissions by the Middle Holocene (ca. 5,000 yrs. BP). Proponents of this theory contend that farming activities, increasing throughout the Holocene, led not only to elevated levels of atmospheric CO₂ but also to rapidly rising levels of atmospheric methane (CH₄) because of the pervasive adoption of rice paddy farming in Asia. Modern livestock production has further contributed to anthropogenic methane emissions via cattle and swine production, CO₂ emissions via fossil fuel consumption, and N₂O emissions via nitrogen-based fertilizer use. While the timing of agriculturally driven increases in greenhouse gas emissions is still a matter of some debate, agriculture today is clearly a major contributor to the anthropogenically enhanced greenhouse effect.

Agricultural Impacts on Water Supplies

The use of irrigation is a hallmark of modern agriculture although the practice goes back for many millennia. The settling of the arid American
West and the rise of the megalopolises of that region were contingent on the impoundment and redistribution of surface waters from the Colorado River. The impoundment and diversion of surface waters for agriculture is one of the reasons why as many as 50% of the world’s freshwater fish are currently threatened with extinction. Agricultural practices in arid regions led to an increase in the salinity of surface waters, and the overpumping of groundwater for agriculture has resulted in subsidence, surface fissures, and saltwater intrusion.

One classic example of the impacts of agriculture on surface water supplies is the case of the Aral Sea in Central Asia, once the fourth largest freshwater lake in the world. Beginning in the 1950s, water was diverted for a large-scale cotton irrigation project. The diversion of the Amu Darya and the Syr Darya, the two largest tributary rivers to the Aral Sea, led to a dramatic reduction of the lake’s surface area, a collapse of its large commercial fisheries, and a dangerous increase in dust storms carrying lake-bed salts and pesticide residues.

The Green Revolution and Modern Farming

Global food production took a dramatic turn during the two decades immediately following World War II. In what was widely perceived as a triumph over Malthusian doomsday food security scenarios, the Green Revolution (GR) ushered in a period of rapidly increasing global food production beginning around 1945 that gained momentum through the 1970s and 1980s. The GR came at a time when the world was struggling to feed its burgeoning populations, especially in the developing countries of Africa, Asia, and Latin America. The positive impacts of the GR are undeniable—famine reduction, food security for developing countries, and an incalculable number of lives saved. The environment also benefited from the GR in that more food was produced on the land under cultivation, potentially saving huge forested areas from imminent conversion to agriculture. In Asia, for example, cereal yields were doubled from 1970 to 1995, while the total land under cultivation rose by just 4%.

The GR, however, has not been such a boon for the overall health of global environmental systems, and it gave rise to several new environmental challenges. As chronicled by Rachel Carson in her book *Silent Spring*, new chemical herbicides and pesticides such as DDT began to have detrimental impacts on nontarget organisms such as the iconic and quintessentially American bald eagle. Increased use of inorganic fertilizers and widespread adoption of new crop varieties have also had negative environmental consequences.

### Pesticides and Nutrient Loading

The GR emphasized inputs of pesticides, herbicides, and inorganic fertilizers. These compounds, however, do not remain on the field surface where they are applied. Pesticides and herbicides can be ingested by nontarget organisms. In some cases, concentrations of these toxic substances increase as they move up the food chain via a process called biomagnification, resulting in increasing toxicity at higher trophic levels. Although many of the most immediately toxic herbicides and pesticides (such as DDT) have been banned or restricted, the longer-term impacts of common agriculture poisons are still not well understood.

The impacts of agriculture on the carbon and methane cycles have already been outlined here, but increased levels of greenhouse gas emissions are not the only changes that agriculture has wrought on global biogeochemical cycles. While the carbon cycle has been clearly altered by agricultural land use practices, perhaps even more striking are the impacts that agriculture has had on the nitrogen and phosphorus cycles. Nitrogen and phosphorus are essential nutrients for plant growth, but they are usually limiting in natural ecosystems. Agricultural applications of these nutrients to field surfaces result in nutrient-rich runoff from the fields. The leakage of nitrogen and phosphorus from field surfaces to rivers, streams, lakes, and oceans leads to the process of eutrophication and in some instances the eventual depletion of oxygen, necessary to support freshwater or marine organisms such as fish. The “dead zone” in the Gulf of Mexico just off the Mississippi River delta is directly attributable to farming in the Mississippi River basin and associated inputs of nitrogen and phosphorus to Gulf waters.

*Robert A. Dull*
ENVIRONMENTAL IMPACTS OF CITIES

A city’s urban environment—the area that it occupies—is directly affected by the city’s use of materials, air, water, and land. A city’s ecological footprint is the space needed to support the city, equivalent to the physical and biological regions disturbed to provide resources to the city and accept its wastes. A city’s metabolism is the sum total of materials, fuels, water, and goods that flow into a city to sustain its populace and economy. Cities vary in size, from agglomerated urban settings of 5,000 people or more to megacities, agglomerated urban settings with more than 10 million people. Global cities are a subset of megacities, cities that set global economic and cultural trends. All cities influence the environment. Large cities tend to have a disproportionately greater environmental impact than less populous ones.

At the turn of the 21st century, cities occupied only 3% of Earth’s land surface; however, their impacts extended across the globe. Cities’ voracious appetites for resources and their dispersal of pollutants into air, into water, and on land result in local, regional, national, and global environmental degradation. For example, in 2000, Los Angeles, California, had a population of about 4 million people and city limits covered an area of 1,290 square kilometers. The Los Angeles metropolitan area had more than 12 million people, including Los Angeles and Orange counties. The greater Los Angeles area, a megacity, covered much of five counties and included more than 17 million people. Direct effects of Los Angeles within the five counties of the megacity are smog, displacement of native habitats, contamination of surface and unconfined ground waters, and deposition of litter on city streets and beaches. The megacity’s direct environmental impacts on the Western United States include air pollutants that diminish visibility at national parks of the Intermountain West, mountains of landfills for urban waste, and diminished flow and increased salinity of the Colorado River. However, the direct impacts to the region are only the beginning of the megacity’s environmental impacts. Coal-fired plants in the west and Midwest provide electricity to Los Angeles, with associated environmental effects of surface and underground mining, coal combustion, and waste disposal. Power plant emissions send sulfur dioxide across national boundaries, affecting Canada’s forests. The impacts of cars driven in Los Angeles are not limited to their emissions. Materials for cars come from mines, chemical plants, manufacturing centers, and cities that dispose of wastes on land, in the air, and into water. Beyond the impacts on local, regional, and global physical and biological environments are the effects on local, regional, and global populaces to sustain the city’s appetite for goods and services. Los Angeles is one of 24,000 urban areas globally. It can be argued that the uneven distribution and unequal per capita
consumption of Earth’s resources exemplified by Los Angeles exacerbate rifts among rich and poor nations and lead to war, famine, and pestilence.

Geographers study the webs of relationships among people, places, and environment. Almost every field of geography addresses an aspect of urban environmental geography, including physical geography, which explains (a) some of the reasons why cities are located where they are, such as regional resources and connectivity to transportation corridors, and (b) the probable consequences of urban development to Earth systems, and human geography, which explains (a) the growth of cities, by migration and natural increase, and (b) the probable consequences of urban development to social and behavioral systems.

### History of Environmental Effects of Cities

In less than a few 100 years, the environmental effects of cities have changed from being largely absorbed by their locale to being globally pervasive and pernicious. Three hundred years ago, the world’s population of approximately 600 million lived in what today would be considered a rural setting with scatterings of villages and towns and about two dozen cities with more than 100,000 inhabitants. Even at their worst, the environmental effects of most cities were largely confined within city limits: local sinkholes of misery, disease, stench, and smoke-filled interiors; outdoor scenes of rotting refuse and human feces; ecosystems supporting rats, vermin, and disease; and social structures vulnerable to recurrent famine.

Cities changed irrevocably in response to the widespread consumption of fossil fuels introduced during the Industrial Revolution and the increased crop yields of the Green Revolution. People migrated to cities in large numbers, leading to increased consumption of resources. Commerce and trade provided jobs, attracting more people to cities. In 1800, only 3% of the world’s 800 million to 900 million people lived in towns of more than 5,000. Fewer than 45 cities had populations greater than 100,000. In 2008, over half of the world’s 6.5 billion people are urban. Cities have evolved from local and regional centers of trade into drivers of national and global economies today.

The history of the environmental effects of cities is inseparable from the history of the environmental effects of the 20th century. The demand for goods and services in cities grew at a faster rate than did their increase in population due to (a) increased per capita consumption, (b) socioeconomic factors such as cultural migrations, and (c) political and economic policies, such as treating the costs to the environment as economic externalities. The metabolism of cities varies among developed versus developing nations and even within nations. The daily intake and output of water, energy, and food for a city of 1 million people in the United States in the 1980s is estimated as more than 500 million kg (kilograms) of water, almost 10 million kilojoules of energy, and almost 2 million kg of food. The daily output included somewhat less than 500 million gallons of sewage, almost 1 million kg of air pollutants, and 10 million kg of solid waste. Calculating and analyzing cities’ metabolisms and their ecological footprint challenge geographers to agree on methodology as well as to model flows of materials into and from cities.

### Environmental Effects on Earth Systems

Every city on every continent has an impact on the Earth. The cumulative environmental effects of cities are innumerable, with feedback loops that make a full assessment of the total environmental impact of a city almost impossible. Assessing even the local effects of cities on their urban environment requires assumptions about costs and benefits, the intangible values of quality of life, and issues of social justice. Although the litany of effects is long, it is possible to get a sense of the diversity of impacts and their scale by listing examples of impacts for each of the subsystems of Earth systems: the geosphere, hydrosphere, atmosphere, biosphere, and anthrosphere.

#### Geosphere

Impacts to the geosphere within the city’s borders include land surface disturbance, such as changes in albedo and hydrologic capacities, and the consequences of waste disposal. It is estimated that globally, one third to two thirds of solid waste of cities is not collected, resulting in
contaminated land and water as well as a stench. Effects of cities on the land within their borders are minor compared with the environmental repercussions beyond their borders. At regional and global scales, the Earth’s geosphere is affected by cities’ needs for food, construction materials, and resources to sustain their economies. Cities’ needs for food result in global losses of topsoil from wind erosion of cultivated fields and water erosion following deforestation. Addition of nutrients, such as phosphate, to increase crop production relies on the mining and processing of chemicals, with attendant environmental effects. Cities’ needs for construction materials result in mining and resource extraction of sand, gravel, limestone, and road materials, including petroleum extracted for asphalt. Cities’ needs for raw materials to sustain their economies require mining for metals, petroleum extraction for plastics, and agricultural disturbances to obtain animal products, in addition to land impacts associated with extraction of coal for electricity and petroleum products for transportation. Humans are geomorphic agents of erosion and deposition. Earth scientists use the term *Anthropocene Epoch* for the time frame of human impacts and to distinguish the sediment record on continental shelves and elsewhere of Earth moving attributable to humans.

### Hydrosphere

Cities need water for their populace to drink. Most cities are located near water supplies. As cities grew, their impacts on local and regional water sources increased beyond the resources within their city limits. Cities in arid and semi-arid regions disrupt natural water regimes by importing water for municipal and industrial uses and withdrawing groundwater from aquifers more rapidly than it is replenished. Agricultural demands of cities justify the construction of dams, diversions, and drainage systems. Cities also use water to dispose of wastes. Until the economic and social consequences of disease outweighed convenience and local economics, sewage and liquid industrial wastes were discharged into local surface waters. It is estimated that even in the first decade of the 21st century, 220 million urban dwellers did not have clean drinking water and 420 million did not have basic latrines. Shallow, unconfined groundwater of most urban areas is biochemically degraded with by-products of human waste, fertilizers, trace metals such as mercury, and emerging contaminants such as pharmaceuticals and antibiotic-resistant pathogens. Contamination of waterways associated with cities includes contamination of oceans and waterways used to transport materials to and from cities and impacts including petroleum spills, flushed bilge water with invasive species, and ubiquitous plastic flotsam.

### Atmosphere

Cities depend on diverse fuels for heat, transportation, and industry. Combustion of fossil fuels emits gases that affect local, regional, and global air quality. Effects within city boundaries differ with topography, culture, and national regulations. It is estimated that in 1998, more than 1 billion city dwellers breathed unhealthy air. Breathing the air pollutants for a day in Kolkata (formerly Calcutta), India, equates to smoking a couple of packs of cigarettes. The impact of cities on the atmosphere is rarely constrained by city borders. Combustion of fossil fuels for transportation within and between cities discharges carbon dioxide and nitrous oxides into the atmosphere. Cities’ needs for electricity and heat, largely met by burning coal and natural gas, release sulfur dioxide into the atmosphere, with downwind consequences of acid rain and increased greenhouse gases such as carbon dioxide and water vapor. Cities’ needs for food result in emissions of methane and particulate pollution from agriculture. Cities’ needs for raw materials for industry result in a potpourri of gaseous pollutants emitted locally and at their places of refining or manufacturing. The cumulative effects of local and global discharges of pollutants into the atmosphere, largely to sustain urban economies, are the inconvenient truths of global warming and climate change.

### Biosphere

Cities affect the biota within their borders and, to an even greater extent, the biosphere of regions that supply their food and raw resources. Urban ecosystems favor generalists, omnivores such as
rats, animals that adapt to buildings such as pigeons, and biota appreciated by humans such as cats, dogs, and grass for lawns. Urban land uses displace natural biota. Urban corridors fragment habitats. Those impacts are minor compared with the effects of agriculture to sustain the growing urban need for food. In 1900, 20 million km² (square kilometers) of Earth was cultivated, and in 1990, 40 million km², the difference largely to feed city dwellers. The oceans, fished for millennia as an apparently inexhaustible source of nutrition, have been harvested of some species of mammals and fish to the brink of extinction. Aquaculture brings environmental challenges analogous to feedlots for cattle. Meanwhile, globalization and increased concentrations of urban populations have accelerated the spread of invasive species and enhanced dispersal of diseases to animals as well as to humans. Even more extensive than a city’s effects on its urban biota or its effects on regional and global ecosystems will be the as yet unknown effects of atmospheric emissions on global climate and, consequently, the impacts on Earth’s biosphere.

**Anthroposphere**

Not the least of the effects of cities on Earth systems are those on the anthroposphere, the sphere of Earth inhabited by humans. Cities are magnets for people, culture, and amenities. They are places of power, of jobs, of promise. However, the voracious appetite of cities for resources has social repercussions within cities and beyond their borders. Large cities are havens of poverty. It is estimated that between 25% and 50% of urban populations live in poverty. Environmental effects of urbanization and globalization have widened the divide between rich and poor within cities and among nations. One international environmental consequence of national materials policies is the exported pollution from rich countries, for example, those that regulated emissions from smelters to poor nations willing to accept the package deal of jobs and degraded air. Nations rich in fossil fuels have a natural advantage over those with minimal resources. Social justice competes with economic realities. Economic health competes with human health. Nations wage wars. People in cities feel the pain but also cause it.

**Future Options**

The environmental effects of cities are pervasive, persistent, and population driven. They are not inexorable. In fact, it can be argued that the environmental impacts of the populace concentrated in cities should be less than those of the populace spread more evenly across Earth’s surface. Cities’ environmental problems challenge modern society’s mores because recent advances in the quality of human life have been tied to per capita consumption, gross national product, and fuel consumption, all of which have a negative impact on the environment. Cities can and have reversed the environmental degradation of their air, water, and land. The “environmental miracle” of Japan during 1965 to 1985 not only cleaned up the fouled air and water but also strengthened Japan’s economy, establishing its reputation as the premier manufacturer of pollution control equipment. Collaborations of government, industry, and individuals, including scientific elites, has reduced acid rain from coal-fired plants in Europe, banned CFC (chlorofluorocarbons), and reversed the eutrophication of the Great Lakes.

**Geography Matters**

Environmental issues are issues of human geography as well as physical geography. The study of geography provides ways to appreciate the differences among cities and their impacts. Urban environmental geography examines how physical geography influences the viability and vulnerability of urban settings. Human geography helps explain why migrations into and among cities has had a greater impact on the environment than has an equivalent growth by natural increase.

The 20th century at first was ignorant of, and then denied, the environmental effects of cities, relegating debates to the concerned scientists and environmental activists. Disparities between the rich and poor nations driven by consumption now affect the world balance of power and undermine world stability. Issues of environmental and social justice include the dynamics of international aid, consequences of famines, and politics of population policies. In 2007, former U.S. Vice President Al Gore and the UN International Panel
on Climate Change received the Nobel Peace Prize for their work in identifying, quantifying, and publicizing the effects and probable consequences of human impacts on the environment. That the prize was given for peace is a signal that the 21st century may recognize that the environmental effects of cities extend beyond city boundaries and beyond their global ecological footprint to the social welfare of the human species.

Genevieve Atwood

See also Ambient Air Quality; Anthropogenic Climate Change; Atmospheric Pollution; Climate Change; Coastal Zone and Marine Pollution; Environmental History; Environmental Impacts of Manufacturing; Environmental Impacts of Roads; Environmental Planning; Environmental Protection; Environment and Development; Greenbelts; Green Building; Green Design and Development; Human Dimensions of Global Environmental Change; Landscape Ecology; Love Canal; Nature-Society Theory; Photochemical Smog; Population and Land Use; Population, Environment, and Development; Public Water Services; Recycling of Municipal Solid Waste; Regional Environmental Planning; Resource Economics; Social Construction of Nature; United Nations Environmental Summits; Urban Ecology; Urban Environmental Studies; Urban Gardens; Urban Geography; Urban Green Space; Urban Heat Island; Urbanization; Urban Land Use; Urban Metabolism; Urban Solid Waste Management; Urban Sprawl; Urban Storm Water Management; Urban Sustainability; Urban Water Supply; Waste Incineration; Water Needs; Water Pollution; Zoning

Further Readings


Energy

The post–Industrial Revolution use of coal as a fuel seriously affected the environment at extraction and usage sites. Soot and sulfuric and hydrochloric acid fumes were an omnipresent bane of urban life in the 19th century. Using electricity rather than burning coal or coal gas directly and the switch to smokeless fuels improved air quality. However, advances in science and medicine
transferred environmental concern to microscale particles and acidic aerosols. Attention shifted from local to continental, and finally global, impacts, resulting in international negotiations over air quality protection (e.g., the 1979 United Nations Economic Commission for Europe Convention on Long-Range Transboundary Air Pollution and the 1998 Kyoto Protocol to the United Nations Framework Convention on Climate Change). Energy-intensive industries, including electricity generation from fossil fuels, are the main contributors to air pollution and greenhouse gases such as carbon dioxide (CO₂). The major industrial air pollutants are acidic aerosols (e.g., sulfur, nitrogen oxides), volatile organic compounds (including contributors to ground-level ozone), particulates (especially those less than 10 micrometers in diameter, known as PMₜ₀), and toxic metals (e.g., arsenic, chromium, lead). Although air pollution control technology can keep emissions below the prescribed levels, it generates potentially hazardous solid or liquid waste products.

**Water**

Water courses and coasts provide economical means to import raw materials and export goods, as well as sources of water for production and a sink for discharged effluent. Groundwater is also at risk of impact. Excessive water abstraction can change the properties of a river, altering biodiversity and the landscape. It can also increase pollutant impact by decreasing dilution. Water returned to rivers can cause damage if insufficiently cooled. Potential surface and groundwater pollutants from manufacturing effluents include volatile organic carbons (e.g., petroleum products and industrial solvents), detergents, heavy metals, and food production by-products. The environmental protection of water bodies and surrounding environments has radically improved in industrialized countries over the past 30 years with the advent of emissions and conservation regulations, though the out-migration or closure of many industrial plants has contributed to this.

**Materials**

Raw material extraction and processing have environmental consequences that occasionally can be restorable to an extent. These consequences may be remote from the manufacturing location. Waste is not synonymous with pollution, but it may imply an inefficient use of raw materials and requires treatment and/or disposal. Some industrial waste or by-products can be fed back into production; others can become “raw materials” in complementary manufacturing sectors. Products may have an environmental impact during use. German, European Union, and other producer responsibility regulations extend manufacturers’ and importers’ liability to include products’ end-of-life phase, to both divert waste from landfills and encourage eco-design. Extraction and processing of raw materials and the disposal of by-products and products each requires a transportation infrastructure (water, rail, road, and/or air based) up to global in scale.

**Geography of Impacts**

The environmental impact of manufacturing has increased from local to global with the advances in industrial technology and increasing production scale. However, the impact is not uniform. Regulatory and voluntary initiatives in industrialized countries help reduce local-scale impacts; the air and rivers in industrialized nations are cleaner than they have been since the Industrial Revolution, although energy usage and waste production have increased. Even so, improvements in industrialized countries in significant part reflect the shift of manufacturing to Asia and Latin America. The decline of heavy industries in Europe and North America through deindustrialization revealed a legacy of contaminated land. Decades of industrial use with limited regulatory restraint resulted in contamination, for example, with phenols from industrial solvents and heavy metals such as lead and arsenic. Concern has also risen over “nuisance” pollution, for example, odor, noise, and traffic. Conversely, industrializing countries face the challenges of poverty and underdevelopment alongside the environmental impacts of industrial production and, increasingly, consumption. Environmental regulations may not be accompanied by the requisite funding, experience, or ability to enforce them.

While local- to continental-scale environmental impacts have shifted rather than disappeared,
environmental concern in industrialized countries focuses on the global-scale issue of CO\textsubscript{2} emissions to the atmosphere. In addition, the potential environmental impacts of emerging biotechnology- and nanotechnology-based industries are not yet well understood.

**Theorizing Environmental Protection**

There are several theoretical approaches to understanding the environmental impact of manufacturing, especially the extent to which it can be contained. The separation of economic growth from environmental impacts is a major goal of environmental policy, though improvements to efficiency (decrease in pollution emitted per unit of production) can be offset by production increases. As resource use and waste disposal represent costs for manufacturers, environmental efficiencies can bring economic benefits—the win-win of eco-efficiencies. Some argue that, conversely, the capitalist imperative for growth inevitably erodes environmental value.

*Pauline Deutz*

See also Anthropogenic Climate Change; Atmospheric Pollution; Bhopal, India, Chemical Disaster; Brownfields; Chemical Spills, Environment, and Society; Chlorofluorocarbons; Climate Policy; Coal; Deindustrialization; Ecological Modernization; Energy and Human Ecology; Environmental Impacts of Cities; Environmental Protection; Industrial Ecology; Industrialization; Industrial Revolution; Open-Pit Mining; Petroleum; Polychlorinated Biphenyls (PCBs); Water Pollution

---

**ENVIRONMENTAL IMPACTS OF MINING**

See Open-Pit Mining

**ENVIRONMENTAL IMPACTS OF OIL FIELDS**

Development of oil fields—sites containing oil and gas reserves—may cause various changes to the natural and human environment of large geographic areas. The environmental impacts arise during the offshore or onshore oil field operation phases, which are exploration, development, production, and transportation. Environmental impacts may be direct if caused by the actual oil field operations, for example, water contamination by drilling wastes, or indirect if they result from other activities arising from oil field developments, for example, job-related migration. An overview of basic direct environmental impacts of oil field operations is essential to understand the nature, scope, and gravity of their influence on both biophysical and sociocultural environments.

Oil field development starts with broadscale geological and geophysical surveys to identify exploration targets. Offshore exploration usually consists of aerial geosurveys, ship traverses, bottom sampling, seismic surveys with explosives, and test drilling for geological data. These activities may result in disturbance of marine mammals, benthic communities, coral reefs, seafloor, and sediments as well as increase in water turbidity. Onshore seismic exploration may trigger environmental impacts such as noise, vibration, air emissions, and intensive surface and ground disturbance from explosive charges, transport traffic, access roads, and drilling of exploratory wells. Whether offshore or onshore, exploration may lead to the degradation of air, water, and soil quality by discharges of drilling fluids contaminated with mud, formation water, and oil. Despite being a nonintense land use with minor effects on existing activities, the exploration phase

Further Readings


may still cause land use conflicts, for example, interfering with fishing and boat use in coastal areas. It may also carry a risk of negatively affecting native communities or damaging cultural resources and sensitive ecosystems.

The onshore development phase consists of construction of roads, airfields, well sites, and production facilities that might modify topography through erosion, sedimentation, and compaction and cause vegetation loss/damage. These impacts may further reduce wildlife habitats, timber yield, and so on, and harm sensitive ecological areas (e.g., wetlands or endangered species). Air pollution might result from wind erosion and soil disturbance due to construction and traffic as well as from burning of fuel fossils that emit CO (carbon monoxide) and NO$_2$ (nitrogen oxides), which in high concentrations are harmful to health. Offshore, the seafloor may be disturbed by anchor dragging, drill ship and platform siting, production facility installation, and pipeline trenching. This alongside the possible discharges of drill mud, cuttings, disturbed bottom sediments, and solid wastes may contribute to bottom contamination and increased mortality of benthic and coral communities. Offshore and onshore installations are commonly accompanied by noise and vibrations, potentially harmful for humans and buildings close to construction sites. Installed facilities may affect natural and cultural landscapes, thereby producing different-scale visual impacts and aesthetic damage. The development phase may accidentally damage cultural resources and historical or religious sites. Labor immigration may overtax community services and cause economic, social, or cultural conflicts.

An offshore production complex includes various types of platforms with production and injection wells, transport pipelines, storage, and primary processing facilities. Platforms are self-contained facilities with helipads, accommodation for work crews, power supplies, storage tanks, and so on. Production requires an extensive shore-based support system for the permanent housing of workers, supplies, waste disposal, and refining. Platforms and drill ships are supplied by both ship and air transport. Extracted oil and gas is transported to shore-based refineries/gas processing facilities by pipelines, tankers, or barges. Air emissions may result from fuel combustion, diesel-powered generators and pumps, barge or tanker product evaporation, operation of oil refineries, vessel up- and offloading, and accidents, for example, evaporation of oil spills or gas discharges from pipeline ruptures or ship collisions. Routine production may cause long-term, minor chemical contamination of the water around platforms.

Onshore production facilities are less complex than offshore ones. They include wells and pumps spaced over the field, gathering and transportation lines, storage tanks, and primary processing units. As in offshore oil field development, here air pollution may continuously result from equipment operations, flaring of unwanted gases, discharge of sour gas and treated formation waters, oil spills, and burning of oil waste.

Emissions of NO$_2$ and SO$_2$ (sulfur dioxide) from oil and gas burning at offshore/onshore production facilities can contribute to large-scale negative environmental impacts (e.g., the greenhouse effect and acidification). NO$_2$ and SO$_2$ can be converted into sulfuric and nitric acids in the atmosphere and deposited on the ground as dry particles after precipitation far away from the emission points, leading to acidification of soil, vegetation, and fresh water and consequently to adverse affects on human health and ecosystems. NO$_2$, CO, soot/dust, and heavy metals released from burning of fossil fuels for transportation and heating may cause serious health problems. Operational noise is generated periodically by gas flaring and regularly by turbines, burners, valves, pipes, and steam outlets.

During any oil field operation stage, catastrophic events may happen, including well blowouts with fire or release of sour gas, platform collapses, tanker collisions and spills, refinery or gas processing plant fires, pipeline breaks, and breakdown of equipment due to corrosion, worn-out components or physical damage, poor maintenance, and leaky oil/water drainage systems. Accidents cause severe soil, air, and water contamination, calling for urgent cleanup activities and leading to additional environmental disturbances. Injuries and deaths from accidents in facility operation as well as increased maritime accidents in coastal areas may occur.
In many cases, environmental impacts from oil field developments can be mitigated or prevented by preproject environmental assessment and consequent environmental management/monitoring plans. The risk of severe air, water, and soil pollution, biodiversity loss, and accidents may be reduced through better technological solutions, energy efficiency measures, compliance with operation, maintenance, and safety guidelines, and development of skills and competences.

Maia Gachechiladze

See also Chemical Spills, Environment, and Society; Coastal Zone and Marine Pollution; Environmental Impact Assessment; Environmental Impacts of Pipelines; Oil Spills; Petroleum

Further Readings


Environmental Impacts of Pipelines

Pipelines are present worldwide, both below and on the Earth’s surface. And as this is an economical and relatively safe transportation method, pipelines are extensively used to transport any type of chemically stable substance, in either gas or liquid form, over land and sea. Some pipelines transport liquids or gases that are related to our living environment, such as drinking water or sewage. Others are used for transport of hazardous substances such as hydrocarbons or chemical wastes from production processes. Because of the toxic and hazardous nature of many of the transported substances, pipeline leaks result in both economic and environmental damage and may even cause human casualties.

Leaks are often difficult to detect when a pipeline is subterranean. When a leak is large or undiscovered for a longer period of time, substantial volumes of gases or liquids may leak into the soil and lead to the development of dangerous situations. Often, leaks also involve a costly remediation. The U.S. National Transportation Safety Board has reported that there have been millions of dollars in losses and several casualties due to gas pipeline leaks in recent years. Depending on whether the pipeline is above the surface or subterranean, the type of material transported through the pipeline, and the location where a pipeline might leak, the impact on the environment will be variable, detection methods might change, and remediation efforts will be different.

Different Pipelines

Different types of pipelines can be distinguished depending on the transported substance, the transport function, the material from which the pipeline is made, and whether it is above or below the surface. Each type of pipeline has its own specific potential environmental problems. Long-distance transportation between cities, countries, or even continents requires major high-pressure pipelines. For gathering and distribution functions, smaller and shorter pipelines are usually sufficient. Major transport lines are especially sensitive to damage since they run through communal areas that are not controlled by the owner of the pipeline and information on the exact location of the pipeline is not always readily available to people living or working in the area. Drilling activities and mechanical ground movement by third parties therefore occasionally result in leaks. In addition, product theft, which particularly happens in developing countries (e.g., Nigeria), leads to problems for pipeline owners, and occasionally casualties. Due to the size of pipelines and the hourly flux of substance through them, large quantities of environmentally damaging substances may be quickly spilled, causing a major environmental impact.
Detection of Leaks

Since oil and gas pipelines are an important economic asset but may cause extensive environmental damage when they leak, regulations on either government or company policies have to ensure the safety of the environment, capital assets, and human welfare. Depending on the type of pipeline, different technologies and strategies have been developed. These range from visual inspection in the field to remote sensing techniques. Traditional inspection methods, such as visual inspection, drilling, and geochemical analysis, are both destructive to the soil and pipeline bedding and time-consuming and thus expensive. These techniques can only be used to investigate the extent of an already existing leak but are too expensive to be used for frequent monitoring. The most common technology for both surface and subsurface pipelines is pipeline integrity monitoring. A pipeline integrity system takes information from sensors in and outside the pipeline to detect anomalies in pressure, flow, stress, and temperature that may result from leakage.

Most types of pollution can also be detected by geophysical means. Ground penetrating radar (GPR) and resistivity/conductivity measurements have been applied widely in environmental studies. An advantage of these techniques is that they provide information on a larger scale at lower costs. A disadvantage of these methods is the nonuniqueness of the derived signal with respect to the pollution problem. Limited validation through geochemical analysis of drill cores is often needed.

A developing technique in pipeline investigation is remote sensing. Remote sensing offers a nondestructive investigation method, covers large areas in a relatively short time, and has significant added value to traditional methods. It is widely applicable since it offers a method for both direct and indirect detection of the presence of pollutants in the near-subsurface region (the latter through vegetation stress and mineralogical changes in soil composition). In addition, information is provided on the spatial extent of vegetation and soil composition anomalies on the Earth’s surface, which may improve insights into

Figure 1  Substances leaking from pipelines affect the environment, both underneath and above the surface.

Source: Faculty of Geo-Information Science and Earth Observation, University of Twente.
Figure 2  Schematic representation of the role of different techniques in pipeline inspection and monitoring

Source: Faculty of Geo-Information Science and Earth Observation, University of Twente.
pollution diffusion mechanisms (Figure 2). A drawback of optical remote sensing is that the observed anomalies are not unique indicators of pollution.

**Monitoring and Remediation**

Treatment of pollutants from pipeline leaks can be done in situ or ex situ. In situ treatment means that the pollutants are treated at the site, while ex situ involves the removal of contaminated material to be treated elsewhere. Remediation of contaminated soils can be done by chemical and biological treatments. Geochemical treatment methods are adsorption, absorption, filtration, distillation, and immobilization. Biological remediation is done through uptake of pollutants by vegetation (phytoremediation) and fungi (mycoremediation) or by stimulation of microorganisms that break down pollutants. These methods can usually be applied without the need to excavate the contaminant material and dispose of it elsewhere. Not all pollutants are absorbed or captured by plants or organisms, and the pollutants should be prevented from ending up in the human food chain.

Mark van der Meijde and Harald van der Werff

See also Chemical Spills, Environment, and Society; Coastal Zone and Marine Pollution; Ecological Risk Analysis; Environmental Impact Assessment; Environmental Impacts of Oil Fields; Environmental Management; Oil Spills; Petroleum

**Further Readings**


Roads affect physical, chemical, and biological environmental systems in various ways. This entry describes the impacts of road construction, use, and maintenance on both living and nonliving systems. It then discusses the role of road ecology, an applied science oriented toward the mitigation of the environmental impacts of roads.

Roads are networked systems that include the roadway itself as well as the infrastructure required to maintain it, such as shoulders, retaining walls, bridge piers and abutments, and storm water management structures. Road networks provide for the transportation of goods and people, facilitating economic development through the exchange of resources and ideas. Traffic is an integral part of road system function. The environmental impacts of roads can be attributed to both their systemic structure and their function, and they are generally studied at local to regional scales (i.e., 10⁴–10⁹ square meters). Road construction creates significant environmental impacts, including habitat destruction and degradation, soil erosion, and the sedimentation of waterways. The persistent use and maintenance of expansive road systems produce extensive, unremitting environmental impacts. Road systems affect abiotic landscape processes, among them hydrospHERic, atmospheric, and biogeochemical cycles. Climate and soil type are important factors that influence these effects. Roads and their traffic also affect living systems, including animal and plant populations. There are also indirect environmental effects of roads that occur as human activities respond to changes in road network structure or function, usually driven by economic factors.

**Impacts of Roads on Nonliving Systems**

Road systems affect hydrospHERic processes by altering the flow of water on and below the surface of the land. Roads crossing low-lying areas can interrupt the movement of surface water, affecting the available water on both sides of the road. In rivers, bridge piers and abutments can cause changes to stream geomorphology, erosion
and sedimentation patterns, and local aquatic habitat conditions. Drainage swales associated with roads can increase stream density, hastening the removal of water from the landscape, reducing percolation, and increasing peak flows during storm events. Debris flows and landslides occur more frequently in sloping terrain with roads present, especially in tropical environments.

Roads affect atmospheric systems by creating microclimates near roads, changing the airflow and relative humidity, often making the air in the road environment more turbulent, hotter, and drier than in surrounding areas. Other atmospheric effects near roads include higher levels of dust, increased noise, and greater nighttime illumination from both vehicle headlights and roadway lighting.

Roads also affect other natural processes such as fire and biogeochemical cycles. While roadsides are often ignition points for vehicle-related fires, roads also limit the spread of fire and are used to manage prescriptive fires. Biogeochemical processes are affected by the chemical discharge related to road systems and traffic. These include emissions from vehicles, compounds used in maintenance activities, and leachates from infrastructure. The products consist of a wide array of nutrients, greenhouse gases, and contaminants, from heavy metals to hydrocarbons and carcinogenic compounds. Particulate and volatile elements and compounds enter the environment through storm water runoff and atmospheric deposition.

### Impacts of Roads on Living Systems

Roads and their traffic affect living systems in various ways: They destroy and fragment landscapes, affect mortality rates and behavior in animals, and introduce exotic species. Roadside environments, with higher levels of light, nutrients, and storm water runoff, provide opportunities for some species to thrive, creating linear “edge” ecosystems that extend undisrupted across the landscape. Some of these are exotic invasive species, which compete with native species and disrupt ecological relationships. For some animals, roads provide opportunities for movement and are sources of forage or carrion. However, for many amphibians, reptiles, birds, and mammals, roads act as barriers thwarting their movement. Roads sever contiguous land areas, causing fragmentation of habitats. Roads can isolate organisms from one another, undermining the ability of populations to recover from disturbances and local extinction events. Changing traffic patterns create ephemeral disturbances that enhance the barrier traits of a road. Mortality related to vehicle-animal collisions (i.e., “roadkills”) is significant, especially for large mammals. Traffic noise causes behavior modification in several species. For some, even the presence of small, remote roads is sufficient to affect their behavior and survival.

### Road Ecology

The emergence of road ecology marks the coalescence of decades of research on the ecological effects of roads and the efforts of applied scientists and professional planners to grapple with the far-reaching and significant effects of road systems. During the latter half of the 20th century, the road network in the contiguous United States expanded considerably, and most places are now within about 1 kilometer of a road. Amid growing evidence of the environmental impacts of road systems, landscape ecologists and environmental planners proposed “road ecology.” Road ecology is an applied interdisciplinary science that addresses questions about various ecological effects of roads. Efforts to coalesce this incipient field culminated in 2003 with the publication of *Road Ecology: Science and Solutions*. Road ecology theory emerges from the collaborative experience and knowledge of physical scientists exploring broadscale ecological questions and social scientists involved in transportation research. Applications of road ecology thus far tend to focus on mitigating the environmental impacts of roads, including, for example, the design of roadside vegetation planting and maintenance regimes and the development of decision support models for transportation planning.

*Alisa W. Coffin*

See also Environmental Impact Assessment; Environmental Impacts of Pipelines; Landscape and Wildlife Conservation; Landscape Ecology; Transportation Geography; Urban Ecology
Environmental Impacts of Tourism

Tourism is defined by the United Nations World Tourism Organization as the activities of persons traveling to and staying in places outside their usual environment for not more than one consecutive year for leisure, business, and other purposes. Tourism can also be viewed as a form of development that depends on natural resources and the built environment. The environment is at the core of tourism because tourism depends on maintaining the scenic attractiveness of a destination that people want to see and experience. However, as soon as tourism begins to occur, it changes the environment to facilitate its own development.

At one time, tourism was seen as a clean industry, causing few environmental impacts. Although it is the world's largest industry by employment and is acknowledged to produce significant environmental impacts, tourism is not as detrimental to the environment as other large industries such as mining and manufacturing. It was not until the late 1960s that the impacts of tourism on the environment began to be recognized. By 1970, the first studies of the environmental impacts of tourism were made. Today, it is widely acknowledged that tourism produces environmental impacts across time and scale. As such, tourism's impacts are difficult to categorize, quantify, measure, and differentiate from those of other activities or natural processes. Tourism's environmental impacts are also often cumulative, coming from many small sources rather than a few large ones. These impacts can be manifest in the form of solid and liquid wastes; air, visual, and noise pollution; congestion; crowding; soil erosion; and deforestation. The intensity, location, and diffusion of these impacts change over time because of factors such as economics, technology, and social and political changes, making assessment difficult and constant monitoring important to maintain environmental integrity. Tourism also leads to further development of destinations due to land development or through infrastructure improvements such as road repair, increased water supply, and increased capacity to handle waste, attracting other economic activities that might otherwise not have come to the destination. Tourism's environmental impacts also originate in various economic sectors and in activities such as hiking, which are not recorded by the economy. Certain tourist activities are also activities of the destination population (entertainment, shopping), and impacts in a tourism-related sector can cause impacts in other unrelated sectors. All these factors make it difficult to determine the exact environmental impacts of tourism. In addition, the type and intensity of impacts depend on the type of development, the characteristics of tourists, and the characteristics of the destination.

To identify the environmental impacts of tourism, an assessment must be made of the physical impacts of tourism development independently of other activities, the baseline environmental conditions at the destination, the types of flora and fauna, and the ecosystem's ability to handle impacts caused by tourism. Often, the environmental impacts of tourism can be placed into one of three categories.

Direct effects are those that originate directly from tourism activities. These include

- sewage disposal,
- increased pollution from tourist transport (e.g., more solid wastes),

**Further Readings**

Indirect effects on the environment from tourism development are those that occur in a manner not directly related to tourism activity. For example, an increase in tourism may lead to more hotels being built. Builders must purchase more building materials and transport these materials to the site. Thus, extracting and transporting the natural resources needed to build the hotel have indirect impacts on the environment.

Induced effects are results of the overall development brought about by tourism that occur in a manner unrelated to tourism, either directly or indirectly. For example, development of tourism in a region may lead to infrastructure improvements to accommodate more tourists; these improvements also attract other industries, and the environmental impacts of these new industries would be classified as induced effects. Similarly, when regional employment increases due to tourism, the additional spending by the workers leads to further impacts.

These environmental impacts are manifested at different scales. At the local level, there are impacts such as the destruction of vegetation from hiking, trekking, or mountain biking. Regionally, the development of services for tourists (e.g., restaurants and hotels) and the development of recreational attractions (e.g., golf courses and theme parks) have environmental impacts. At the global scale, the burning of fossil fuels in different sectors of the tourism industry contributes to climate change through the release of greenhouse gases (GHGs). Many of these impacts can be attributed to the transportation and accommodation sectors of the tourism industry. In the transportation sector, travel by private
car or airplane contributes significantly to GHG emissions. In the accommodation sector, GHG emissions come primarily from energy used for heating, cooling, cooking, water heating, and lighting.

Because of the complexity of the tourism industry and its interconnections with other economic sectors and industries, it is difficult to quantify environmental impacts. Other factors that make it difficult to quantify environmental/ecological impacts and determine the carrying capacity of a destination are the lack of widely accepted environmental indicators, rapid changes in environmental problems stemming from technology and development, and direct versus indirect impacts, which are often difficult to decipher. Carrying capacity is the most widely used concept in quantifying environmental impacts for tourism planning. In a tourism context, carrying capacity is the ability of the environment to support tourism and associated activities without diminishing the quality of the environment and visitor satisfaction. Carrying capacity is an instrumental concept for guiding environmental management and policy making in tourism. However, in practice, applying this concept has proven difficult. First, it is difficult to identify the critical limiting resource to use as a basis for estimating carrying capacity. Second, the number of people a destination can handle depends on the type and intensity of tourist activity. Third, there are often many explanatory factors for environmental degradation, of which tourism might only be one.

Noting that tourism is a form of development that depends heavily on a healthy environment and that tourism changes the environment to facilitate its own development, the industry response has been to promote sustainable tourism and ecotourism. Unfortunately, there is very little environmental monitoring that occurs to accurately provide land managers with scientific data with which to determine the local carrying capacity and plan ecotourism development accordingly. Even so, both ecotourism and sustainable tourism not only seek to limit the negative impacts of tourism on the environment but also attempt to promote conservation through education, responsible behavior, and environmentally friendly development.

Keith Bosak

See also Carrying Capacity; Ecological Footprint; Ecotourism; Sustainable Development; Tourism

Further Readings


ENVIRONMENTAL IMPACTS OF WAR

War strategy has historically been configured by environmental factors in various ways. Armies have used forests for cover, mountains and valleys for strategic advantage, and rivers and waterways for transport of arms and materials. At times, deliberate destruction of natural resources has also been a military strategy through “scorched earth” campaigns that were practiced as early as the Scythian wars against the Persians around 500 BC to deprive the enemy of food crops. Most recently, Saddam Hussein deliberately set ablaze Kuwaiti oil wells in 1991 to diminish the economic potential of the adversary, leading to widespread air and water pollution in the Persian Gulf region in what was called an act of “ecocide.”

Often, strategic targets that are chosen for bombing have a particularly damaging impact on the environment. For example, the bombing in 1999 by the North Atlantic Treaty Organization of the Serbian city of Pancevo (in the former Yugoslavia) targeted fertilizer-manufacturing sites, chemical plants, and a refinery to reduce the productive capacity of the adversary. However, the most notable impact was the release of pollutants across a wide area of the surrounding countryside.

Apart from the direct and deliberate linkages between ecology and war, there are also specific
indirect impacts of warfare on the environment since resources and physical space are needed to support the military-industrial complex. The ecological footprint of war has increased dramatically over the past century due to the need for massive material and energy sources to build and execute weaponry. Most weapons are still manufactured with metals despite the rise in composite materials for making some aircraft components. The energy consumption for military operations is staggering. For example, the U.S. military consumed 1,100 trillion British thermal units of energy in 2006 (1% of total U.S. energy consumption), which is equivalent to the entire energy consumption of Nigeria. Furthermore, the use of various chemicals, including radioactive materials, is concentrated in military-industrial complexes. As reported to regulatory agencies, the U.S. military on average generates about 750,000 tons of toxic waste material annually, making it the largest polluter in the country (greater than the five largest chemical companies combined).

The physical space needed for research and testing of military weapons is also staggering. Entire island archipelagos have been displaced for military testing practices, such as Bikini Atoll in the South Pacific (site of nuclear weapons tests) and the Chagos Islands in the Indian Ocean (the site of the British military base of Diego Garcia).

A latent environmental impact of war is also evident in policy prioritization. During times of war, environmental factors are often considered
ENVIRONMENTAL IMPACTS OF WAR

secondary to national security priorities, and accountability of armed forces in terms of environmental compliance is greatly diminished. For example, prior to the Iraq War, the U.S. military asked for a series of exemptions from environmental regulatory compliance and was granted several allowances pertaining to endangered species, hazardous waste disposal, and air emissions through the National Defense Authorization Act of 2003.

Occasionally, areas that have been reserved for military purposes to prevent any development can become ecologically more biodiverse in their “wild” state. The area near the Hanford nuclear test site in the Columbia River basin in Washington State is an example in this regard. The demilitarized zone between North and South Korea has been used as an example of how a state of hostility between armed neighbors can sometimes inadvertently benefit biodiversity. While this region is filled with land mines, the absence of development has allowed trees to grow to heights that support bird nests for endangered avian species such as cranes.

Guerrilla warfare affects ecosystems when fighters cut down trees for firewood and engage in poaching of forest animals for bush meat or for coveted commodity items such as ivory that can be smuggled to finance the battles. Such effects have been particularly prevalent in the numerous African civil wars and the conflicts in southeast Asia. Counterinsurgency tactics in such circumstances from conventional forces can also lead to extensive environmental damage, as exemplified by the use of defoliants such as Agent Orange during the Vietnam War.

Given the global concerns about the short-term and long-term impacts of warfare and the military-industrial complex, armed forces worldwide are beginning to consider the environmental footprint of war, often spurred by national regulations or threats of litigation from environmental activists. The U.S. army has now established an Environmental Command to “lead and execute environmental programs and provide environmental expertise that enables army training, operations, acquisition and sustainable military communities.”

There is no comprehensive treaty on environmental protection during times of war, though the Fourth Geneva Convention does stipulate that widespread, long-term, and severe damage to the natural environment can be considered a violation of the agreement. The Environmental Modification Convention is a treaty that could be expanded to cover the larger environmental impacts of war; it has been ratified by most major military powers, including the United States. The United Nations Environment Programme also has a postconflict assessment branch that has been tasked with preparing environmental impact assessments of wars and to conduct some cleanup efforts. However, this unit still depends on mandates from member states on particular cleanup efforts that have to be financed by an ad hoc group of donor countries.

Considering the scale of military operations worldwide, the environmental impact of war is a continuing challenge for the international community.

Saleem H. Ali

See also Military Geography; Military Spending; War, Geography of

Further Readings


An environmental impact statement (EIS) is a public document presenting the results of an environmental impact assessment (EIA). An EIS is an often lengthy and technical document that describes a proposed public or private development or undertaking and reports on its potential environmental effects and prescribed impact management and monitoring programs. Numerous countries have laws requiring the preparation of an EIS whenever an EIA is completed, including Canada and the United States. The EIS is the primary means of communicating the potential environmental impacts of a project to the public and decision makers such that an informed decision on the proposal can be made.

**Origins**

The EIS originated from the U.S. National Environmental Policy Act (NEPA) of 1969, becoming law in 1970, which required that federal agencies conduct an EIA and prepare an EIS indicating, for any development or action with the potential to affect significantly the quality of the human environment, that the agencies have assessed and considered the environmental and socioeconomic consequences of the proposal. Since NEPA, the requirements for the preparation of an EIS have extended to include both public and private proposals and undertakings; however, the requirements for the assessment of private undertakings that do not involve a government agency vary from one EIA system to the next.

Many of the first EISs completed, under NEPA and elsewhere, were highly technical in nature, focused on detailed descriptions of project design, characterized by large inventories of biophysical data about the local environment, and completed after the development had already commenced. By the mid 1970s, many EISs were thousands of pages in length, contained multiple volumes of facts, and were often of limited value to the public and to decision makers. It was not until the mid 1980s, with the introduction of the scoping process in EIA, that EISs started to play a more meaningful role in development decision making—available to the public and decision makers prior to development actions and containing less technical description and more analytical focus on a limited range of environmental effects and their implications. The scoping processes stressed the importance of focusing EIA on a select few priority environmental components, referred to as valued ecosystem components, for the purpose of more efficient and informative impact analysis.

EISs continued to evolve throughout the 1980s and early 1990s in response to several international developments, such as the 1987 report of the World Commission on Environment and Development and the 1992 United Nations Conference on Environment and Development (the Earth Summit), placing increased emphasis on the interrelationships between project actions, environmental change, and the resulting social and economic impacts. In recent years, with the growth of sustainability awareness, EISs have also begun to address the potential positive impacts of development proposals.

**Contents of an EIS**

The content of an EIS varies depending on the regulatory and policy requirements of the EIA system and on the complexity of the proposal. In the United States, for example, the requirements of an EIS are detailed under the 1978 U.S. Council of Environmental Quality Guidelines. In Canada, general guidance is provided under the Canadian Environmental Assessment Act and by the Canadian Environmental Assessment Agency. All EISs must include a description of the proposed action, including a statement of its purpose, and a description of the affected environment that is sufficient to allow an assessment of the potential impacts of the proposed action. An EIS typically consists of six major components:

1. A description of the project, including a statement of the purpose of and need for the proposed action
2. A description of the affected environment
3. Identification of alternatives to the proposed action and alternative means of carrying out the proposed action
4. Analysis of the impacts of the proposed action and its alternatives on the biophysical and human environment, including also direct and induced impacts
ENVIRONMENTAL JUSTICE

5. A statement of significant environmental effects and the strategies proposed for impact management
6. Identification of monitoring needs and programs

Enduring Concerns

An EIS is the primary means of communicating the results of an EIA to decision makers and to the public. EISs are now completed earlier in the development planning and decision-making process, are more integrative of biophysical and human systems, and are more responsive to emerging sustainability issues and concerns. However, many EISs are still largely inaccessible documents—lengthy and written in a language that is not easily understood by the nonengineer or nonscientist. If the technical and scientific language of an EIS is not translated to a lay language that is easily understood, then the objective of the EIS to communicate the potential environmental impacts of a proposal to the public and to decision makers such that an informed decision on the proposal can be made is not achieved.

Bram F. Noble

See also Environmental Impact Assessment; Environmental Management; GIS in Environmental Management; Public Policy, Geography of; Sustainable Development; United Nations Conference on Environment and Development

Further Readings


ENVIRONMENTAL JUSTICE

Environmental justice is both a social movement and a research frame. Rooted in the U.S. civil rights movement of the 1950s and 1960s, the idea of environmental justice gained traction in the United States during the early to mid 1980s. Proponents of environmental justice seek to ameliorate circumstances where people of color (e.g., African Americans, Asians, Latinos) and low-income earners (i.e., people below the poverty line and the working poor) are disproportionately exposed to environmental harms and/or lack access to environmental benefits. Advocates of environmental justice believe that all people—irrespective of race, ethnicity, gender, or class—have the right to live in a clean, healthy, and safe environment and to enjoy equal access to safe and healthy workplaces, schools, and recreation areas and to safe and nutritious food.

This entry considers the main concepts and issues that are central to environmental justice. It begins by reviewing the historical development of the environmental justice movement and then examines what we mean by “injustice” or “inequality.” Next, the various mechanisms behind environmental injustice (i.e., structures, processes, institutions, agencies) are considered. Academic responses to the social movement are discussed, with a focus on geographers’ contributions. Finally, the entry discusses recent developments in both the social movement and academic literature.

It is important to note that environmental justice is related to but different from environmental racism. Environmental justice is also different from ecological justice or justice to nature. Those concepts refer to efforts to protect animals, plants, and vulnerable habitats and recognize the “intrinsic” value of nature. Some commentators have distanced environmental justice from ecological justice, preferring instead to use the term *social justice environmentalism*. In reality, these three concepts have much in common and often overlap.

The Social Movement

When the environmental justice movement began is uncertain. Some commentators suggest that the movement began in the late 1960s. What is clear
is that by the early 1980s the movement had noticeably mobilized around two concerns: (1) getting law makers and decision makers to recognize that so-called minority groups and the poor were disproportionately exposed to environmental hazards and (2) developing effective remedies for these problems.

During the 1980s, groups of concerned citizens in the Southern United States, accompanied by religious leaders and politicians of color, mobilized to help vulnerable people combat toxic landfill facilities, hazardous industries, and other locally unwanted land uses within their communities. The movement has since expanded from a series of disparate grassroots initiatives to a recognizable collective with an international agenda.

Originally, environmental justice activists asserted that mainstream environmental groups such as Greenpeace, Sierra Club, and the Wilderness Society had ignored the plight of socioeconomically marginalized and vulnerable communities, focusing solely on nonhuman species. The movement’s initial purpose was therefore to develop an alternative environmentalism, one based on direct action, social inclusiveness, and human concerns. More recently, many mainstream environmental groups have assisted in environmental justice struggles.

Early examples of successful campaigns include a predominantly African American community’s fight against the siting of a PCB (polychlorinated biphenyls) landfill in Warren County, North Carolina; a battle by Latino mothers of East Los Angeles to prevent the construction of a waste incinerator in their community; and campaigns by Latino farmworkers against poisoning from toxic agricultural chemicals. The early 1980s to the early 1990s was a decade of progress for the movement. Activists garnered increasing public attention and leveraged local, state, and federal action against practices of intentional targeting of socioeconomically marginalized and vulnerable communities for toxic and unwanted land uses. They achieved this through multiple actions including lobbying, political incorporation, forced buyouts and relocations, and lawsuits. By the early 1990s, the movement spanned the United States and encompassed issues as diverse as industrial “cancer alleys,” uranium enrichment facilities, biotechnology labs, brownfield redevelopment, “transit racism,” lead paint in old houses, and radioactive waste sites. With the rapid expansion of the movement came formal organizations, such as the Citizens Clearinghouse for Hazardous Waste, that served to distribute information, enhance communication between groups, and promote collective action.

Major milestones included a U.S. General Accounting Office (GAO) investigation into the siting of hazardous waste landfills and low-income, nonwhite communities, and the subsequent report, which found a strong correlation between landfills and communities of color. This was followed by a report from the United Church of Christ, which corroborated the GAO findings and concluded that communities of color were intentionally targeted as recipients of hazardous waste and unwanted land uses. This period culminated with the passage of the Clean Air Act (1990), the first national summit of environmental justice advocates in Washington, D.C. (1991), and the formation of the U.S. Environmental Protection Agency’s National Environmental Justice Advisory Council (1993). The movement arguably reached its zenith in 1994, when President Bill Clinton signed Executive Order 12989 to combat the unjust actions of government agencies in administering environmental law.

The movement has diversified since the mid 1990s and become more global in its reach. Participants from the United States attended international conferences and symposia on the environment—such as the United Nations Conference on Human Settlements, Habitat II in Istanbul—and formed alliances with people from developing countries. Internationally, people began to notice that while many environmental problems (e.g., ozone depletion, the Chernobyl disaster, climate change) had global impacts, these were disproportionately felt by impoverished and ethno-racially marginalized populations.

Geographers’ Contributions

Environmental justice scholars have worked alongside activists to research environmental inequities and to substantiate claims of injustice. Although some scholars have criticized methods used for investigating environmental inequities, suggesting that inferior methods have led to
spurious findings, many scholars have corroborated claims of injustice made by ordinary citizens. Some scholars have themselves become environmental justice advocates and activists.

Geographers have strongly contributed to the now substantial volume of scholarly literature on environmental (in)justice in its various guises. These geographers have played a major role in identifying where, when, how, and why environmental injustice occurs (e.g., revealing vulnerable populations, identifying park-deprived communities, and showing how “minority” ethno-racial groups are disproportionately incarcerated). Geographers have also developed numerous methods to identify and ameliorate instances of injustice. These range from mapping the connections between vulnerable populations and sources of environmental harm using geographic information systems (GIS), through mobilizing and empowering citizen resistance, to acting as expert witnesses and lending scholarly credibility to citizens’ struggles for justice. Perhaps most important, geographers have shown how many environmental injustices are place-based phenomena. Understanding that regional racial formations work across a range of scales and through a variety of mechanisms to instantiate and perpetuate environmental inequities is a key contribution from geographers. Similarly, geographers have recognized that different conceptions of justice...
shape on-the-ground inequities and determine the efficacy of potential remedies.

Theories of Justice

Theorists generally recognize that there are four basic expressions of equity. These are (1) equitable distribution—all members of society have uniform access to environmental benefits and are identically exposed to environmental harms, regardless of who they are; (2) compensatory equity—environmental benefits are redistributed to the most disadvantaged and vulnerable populations, offsetting inequalities based on class, race, or gender differences; (3) demand distribution—the most vocal members of society are allocated the most resources; and (4) market-based distribution—people who can afford to pay for goods or services get the best access to those goods or services. Knowing how these different versions of justice are mobilized in environmental justice disputes can enable us to understand better how environmental inequities occur.

Mechanics of Injustice/Inequality

The processes that configure environmental injustices are rarely simple. Oftentimes, several imbricated mechanisms are responsible for why particular communities are exposed to environmental harms or why they have limited access to environmental benefits. These mechanisms include intentional targeting, differential law enforcement, market factors, biased decision making, and limited opportunities for public participation. Running through many of these explanations is what the geographer Laura Pulido has termed white privilege. White people typically have better access to resources such as income, education, and political connections by virtue of socioeconomic and ethnic or racial systems of privilege. These privileges enable most whites to secure the most salubrious environments and to escape the worst hazards.

Intentional Targeting

Many case studies have reported that ethno-racial minority groups and the poor have been systematically targeted by noxious industries and hazardous land uses as sites for their operations. Researchers assert that these social groups lack the power to resist the siting of hazardous facilities in their neighborhoods because they are politically disenfranchised or unemployed or have low-paying jobs, may have higher numbers of dependents, may not speak the dominant group’s language, may have low education levels, may lack health insurance, or may be undocumented workers who are at risk of being deported if they speak up.

Differential Enforcement of Environmental Laws

Research has demonstrated that often government agencies take longer to respond to complaints from low-income, nonwhite communities and are less likely to take punitive action against polluters within these communities and that when they take action, the remediation is less stringent than in white and wealthy communities. Furthermore, people of color are less likely to be employed within these agencies, potentially leading to biased decision making. Commentators label such practices as “institutional discrimination.”

Market Mechanisms

Some scholars—from both ends of the political spectrum—have argued that claims of injustice may be difficult to substantiate if we cannot tell who or what came first—the hazardous environments or the vulnerable community. What this suggests is that if the hazard came first, then the outcome is not an injustice. But vulnerability is associated with notions of choice. The less choice that people have, the more vulnerable they are to exploitation and/or harm. Market effects can compound this problem. When people have low-paying jobs or depend on welfare, they may only be able to afford the cheapest housing. This housing may be dilapidated, infested with vermin, contain lead paint or asbestos, or be located near a polluting factory, landfill, waste incinerator, or other potentially harmful sites or on flood-prone or radioactive land. Likewise, even when an environmental hazard preceded the community, that is, existed prior to the arrival of local inhabitants, urban encroachment means that housing may eventually be constructed near that hazard. And because land values will be lower around hazardous sites, vulnerable populations are likely to become disproportionately concentrated nearby—leading to injustice.
Sociodemographic Change

When a community changes from being predominantly white, or even a mixed community, to one constituted predominantly of people of color, there is a significantly greater likelihood that it will be the recipient of noxious or unwanted land uses. This is largely due to the combination of “white flight” (where wealthy, white residents flee inner-city locations in search of suburban refuges, leaving behind those people without means of escape) with institutional discrimination and intentional targeting.

Ineffective or Nonexistent Public Participation

Many environmental justice advocates assert that inequalities are produced through decision-making processes that are not inclusive. The literature is filled with cases of “back-room deals,” decisions being made “behind closed doors,” and polluting facilities that are approved with little or no community input. Without open, accountable, transparent, and inclusive decision-making processes, the interests of vulnerable communities will likely be subordinated to those of powerful corporations and landholders.

Regional Racial Formations

Finally, how ideas of race and ethnicity are mobilized for different ends in different places is a key concern for geographers studying environmental injustice. Geographers have recognized that race and ethnicity are produced as sets of social relations, which act to create and reinforce disadvantage and vulnerability. For example, if decision makers, lending institutions, and enforcement agencies are dominated by wealthier whites, power relations will likely favor their interests over others.

Questions Regarding Methods and Measurement

Geographers have helped refine how we identify and measure instances of environmental injustice, recognizing that environmental justice research requires more than just investigation of patterns of distribution of environmental hazards or historical processes that fostered unjust landscapes. Research must also examine how communities respond to environmental hazards, how coalitions of resistance develop, how government agencies define and address environmental justice issues, how multiple forms of discrimination can lead to environmentally unjust outcomes, and which remedies are effective against various forms of inequality. Many recent advances in environmental justice research have come from alternative methodologies.

Changing Methodologies

Two decades ago the research emphasis was on proving that environmental injustice existed. Researchers favored quantitative methods such as GIS, which allowed them to examine spatial units (e.g., census collection districts, postcodes, neighborhoods) for ethno-racially differentiated patterns of exposure to environmental harm. Geographers and others found that depending on the measure used, injustices manifested in different places or at different scales. Susan Cutter, for example, illustrated how communities affected by hazardous materials transportation varied according to whether the density of the transport network was examined, the unit of measurement was changed (e.g., spills per mile), and population measurements were included (e.g., spills per capita). For a time, critics and proponents of environmental justice fought over issues such as incomplete data sets, inaccurate data, the variables that were included for analysis, and difficulties with spatially related variables that created problems in attributing causality. More recently, researchers have moved from simply identifying problems to examining how affected communities have responded to these problems. This transition reveals a philosophical shift from treating affected communities as passive victims to active agents.

Quantitative analyses have been joined and strengthened through the use of qualitative techniques, typically through case study approaches. They include archival research, oral histories, ethnographies, focus groups, discourse analysis, and performative methods. This shift has partly been driven by the growing recognition that incidents and experiences of injustice are historically and contextually contingent. For example, although Native Americans and Australian Aboriginals...
likely experience similar impacts from the mining and dumping of radioactive materials in their communities, their experiences of colonization, the ways in which racial formations evolved in their respective homelands, and their capacities for, and types of, resistance are likely to be very different. Geographers have thus become more attuned to issues of essentializing environmental injustice.

**Scale**

Geographers have contextualized environmental inequalities and potential remedies for injustice by investigating how scale affects environmental (in)justice. Research has shifted from a singular focus on the community or neighborhood to cover the full range of geographic scales. Incidents of injustice affect individual bodies—expressed through harms such as cancer, asthma, birth defects, or miscarriage. Hazards occur in homes, workplaces, and education venues; sites of injustice include parks, factories, child care centers, farms, offices, and schools. Inequalities are increasingly manifest across scales from neighborhoods to cities and from states to nations. For example, the global electronic waste trade connects villages, cities, and nations in new webs of exchange and inequality. By studying environmental justice issues across various scales, geographers can illuminate the complex processes responsible for environmental inequalities.

**Expanding the Research Agenda**

Geographers have begun to expand the environmental justice research agenda. Issues studied recently include children’s access to parks and playgrounds, the spatial distribution of recreation facilities, access to fresh water in the developing world, the impact of neoliberal policies on urban forest management, how urban landscapes may foster obesity in minority neighborhoods, and the potential impacts of climate change on socioeconomically and ethnically marginalized populations.

Theoretical frames are shifting too as geographers seek out alternative explanations for injustice. Theoretical frames including political ecology, human ecology, cultural landscape, and ecological modernization are expanding our understanding of how inequalities occur, how they are situated within broader socioecological relations, and how problems might be remedied. For instance, geographers are showing how natural disasters such as Hurricane Katrina are produced through socially mediated ecologies, starkly illustrating how people of color and the poor bear a disproportionate burden of exposure to harm. Katrina has illustrated how the impacts of environmental hazards are increasing not only due to biophysical drivers such as climate change but also due to the intersection of sociodemographic, political, and ideological factors that concentrate vulnerable people in hazardous places and produce harmful nature-society relations.

New analytical techniques such as material flows and lifecycle analyses from fields such as industrial ecology show how environmental harms span commodity chains. They enhance our knowledge of hazards and their spatial and temporal impacts, illustrating, for example, how sites of extraction, production, distribution, and disposal (e.g., mines, factories, railways, and landfills) all potentially generate environmental harms, which affect vulnerable populations in diverse locales. These analyses also show how demographic change, land use planning practices, real estate and property markets, local and national politics, systems of global trade, struggles of access to resources, and war mediate the impacts of environmental harms.

Finally, new research on adaptive capacity and local resilience to hazards is teaching us lessons about when communities are successful in ridding themselves of injustice. Such studies offer us a more nuanced appreciation of why particular environmental laws are not enforced, which lawsuits are successful, where grassroots activism has made a difference, where planning codes have redressed inequality, and whether disaster preparedness programs have actually reduced risks.

*Jason Byrne*

See also Cancer, Geography of; Class, Nature and; Differential Vulnerabilities to Hazards; Ecological Justice; Environmental Law; Environmental Racism; Ethnicity and Nature; Industrial Ecology; Justice, Geography of; Landfills; Locally Unwanted Land Uses (LULUs); Nature-Society Theory; Not in My Backyard
Environmental law refers to the array of rules guiding human behaviors that affect the physical and biological world. These rules define what is prohibited, allowed, or required of people in a given situation and the penalties for transgressing those rules. While rules may arise from a wide range of sources from families to governments, this entry outlines the current framework of U.S. environmental law. It ends with a brief overview of the growing trends of criminalization and internationalization in environmental law.

### U.S. Common Law

The first U.S. environmental laws existed within the common law, a system of law that originated in England based on the decisions of judges from previous cases as opposed to laws created by legislative bodies. Within common law, tort law provides remedies for injuries to a person’s body, property, and rights. For example, a pollutant from one person entering another’s property may constitute a violation of the common law rule of trespass, which prohibits invasion of a person’s property. Several potential remedies exist for the injured party, including the costs to restore and rehabilitate the damaged environment.

### U.S. Federal Environmental Law

Beyond common law, federal, state, and local legislatures and government agencies have created a wide array of environmental laws and regulations. The U.S. federal system limits the national government to those powers expressly stated in the U.S. Constitution, with the remaining powers retained by state governments. States have traditionally handled environmental matters under their police powers, the general authority to regulate for citizens’ health and welfare. The federal government has largely developed the legal framework for environmental protection based on its authority to regulate commerce, foreign relations, and federal property. Moreover, federal law is the supreme law of the land, and federal regulation may preempt any conflicting state laws. Nonetheless,

### Further Readings


state and local governments continue to play an important role in creating and implementing environmental laws and regulations.

Environmental laws generally address specific problems such as water pollution or hazardous waste; however, one major modern environmental law applies across all environmental issues. The National Environmental Policy Act of 1970 (NEPA) requires identification of environmental, social, and economic effects and alternative courses of action for all federal activities that could “significantly affect” the human environment. This occurs through the creation of environmental impact statements (EIS). Within this process, government agencies elicit and respond to public comments before making a final decision. The record of the final decision must include justification for the decision, but it only requires consideration of alternatives and the public’s comments. Thus, NEPA does not guarantee any protection for the environment, only that environmental effects are considered in government decision making.

Pollution control is the prototypical environmental concern and is exemplified by the Clean Air and Clean Water Acts. The Clean Air Act (CAA) requires the U.S. Environmental Protection Agency (EPA) to set National Ambient Air Quality Standards (NAAQS)—concentration limits for “criteria” pollutants, those likely to pose a danger to human health and welfare. States must then devise implementation plans to meet the NAAQS, with EPA providing final approval of those plans. Airsheds on both ends of the spectrum, those not meeting NAAQS and those with high air quality, generally require stricter emission standards or particular technologies for new or modified sources to either improve or protect air quality in those geographic areas. Additionally, the federal government sets technology standards for pollutant sources not covered under the state implementation plans, hazardous pollutant sources, and mobile sources, including automobiles.

The Federal Water Pollution Control Act, commonly known as the Clean Water Act (CWA), aimed to eliminate water pollution discharges by 1985. While water pollution has decreased, it still exists, and the CWA provides the continuing legal authority for federal management of water pollution. The CWA approach involves technology requirements for “point sources,” which are discharges via a “discrete conveyance” such as pipes and ditches. The EPA provides grants to projects aimed at reducing nonpoint sources, for example, polluted snow or rain runoff from agricultural operations or parking lots. Additionally, states must set water quality standards and total maximum daily loads for pollutants in specific water bodies.

Two principal laws regulate hazardous wastes. First, the Resource Conservation and Recovery Act (RCRA) created a “cradle-to-grave” approach that governs the generation, transportation, treatment, storage, and disposal of hazardous wastes. Second, the Comprehensive Environmental Response, Compensation and Liability Act (CERCLA), commonly known as Superfund, funds federal government cleanup of hazardous waste sites, releases, and spills. Superfund also assigns responsibility for cleanup or the costs of cleanup to almost any party even remotely involved with the site, release, or spill, even if their actions occurred before the passage of the Superfund legislation. Additional laws governing toxic substance include the Toxic Substances Control Act, Residential Lead-based Paint Act, and the Federal Insecticide, Fungicide and Rodenticide Act.

Several laws govern the use of natural resources including energy production, water usage, public lands, timber harvest, mining, fishing and hunting, zoning, and soil conservation. Within this body of natural resource issues, two areas are frequently included within the ambit of environmental law, wetlands and endangered species protection. The CWA requires a permit for any discharge into wetlands and as implemented helps prevent many forms of potential wetland degradation, including draining, or at least requires the creation or repair of wetlands to offset any losses. The Endangered Species Act aims to halt the extinction of species and rebuild endangered and threatened species populations, mainly by prohibiting anyone from harming, harassing, or killing listed species.

Criminalization is playing an increasingly important role in the realm of environmental
law. The law is divided between civil and criminal codes, with the type and severity of penalties (e.g., imprisonment for crimes) constituting the main differences between the two. Many environmental laws include criminal penalties, and governments are increasingly prosecuting polluting activities as crimes. Additionally, criminal prosecutions for activities related to the degradation of the environment have relied on laws not directly pertaining to the environment, such as the organized crime–focused Racketeer Influenced and Corrupt Organizations Act and securities laws.

Further Readings


Jeffrey W. Henquinet

See also Energy Policy; Environmental Impact Statement; Intergovernmental Environmental Organizations and Initiatives; Law, Geography of; Neoliberal Environmental Policy; Regional Environmental Planning; United Nations Environmental Summits; United Nations Environment Programme (UNEP)
distinct functions and structures across government agencies and community groups. In asserting control over specific human-environment relations, these dominant discourses have created an artificial separation of humans from the “natural” environment. Environmental management is more than simply using science to classify and make sense of the world in order to assert human control as management. Recognizing the connections and inherent relationships among species provides windows for more ethical and contextualized engagements with place. Reorienting and balancing environmental management discourses to recognize the importance of process and connection provides more robust and appropriate approaches to ensuring socially just ecological sustainability.

This entry contrasts this dominant “managing the environment” approach with approaches to human-environment relations that highlight the connectivity and relationships between all species, including humans. In particular, it balances the discourses of environmental management by providing space for marginalized approaches, such as indigenous understandings, which focus on the connections between people and place and the interrelationships between all species. It advocates acknowledgment of and commitment to diversity in understanding and managing human-environment relations in the future.

**Managing the Environment**

Most dominant discourses of environmental management focus on the need for humans to manage the environments in which they live. Western natural science is often the unquestioned basis for such management decisions and is considered the most important and appropriate mechanism for understanding humans’ relationships with the environment. Many textbooks, for example, argue that science should be the sole approach for making environmental management decisions. Scientific discourses typically allow for unambiguous categorization of the environment into its components, as if those categories are self-evident and universal—an approach that is often displayed by government environmental agencies, which separate, for example, the functions of water management from air quality, land from sea, and conservation areas from urban spaces. Such a perspective denies the intimate connections among these various environmental domains.

When they are viewed from a scientific perspective that is assumed to be objective and value free, environmental issues and resources can be quantified, understood, and thus managed. Water can be allocated (or overallocated in some areas), commercial species can be harvested to sustainable levels, carbon emissions can be measured and traded, and areas of high biodiversity can be “conserved” in a so-called natural state. Environmental management systems deliver quantified and objective targets, such as the international standard ISO 14000 of the International Organization for Standardization, which establishes targets and milestones for organizations to minimize their environmental impacts. Such approaches have been important for reforming and restructuring modernist resource management regimes.

Eurocentric science was developed on the assumption that experiments that could repeatedly produce the same results provide an understanding of nature regardless of the specificity of context and independent of local place effects (i.e., universal laws). In short, most natural science has treated local geographic context as only marginally relevant to the application of broad theories of how nature works. In developing taxonomic systems, in which the world could be understood and classified, scientific discourses also represented humans as separate from and superior to their environment, in the process justifying management intervention and control. Various elements of the environment, such as forests, wetlands, fisheries, or water bodies, which in reality only exist in intimate connection to one another, could thus be theoretically separated from their context and from one another and managed as discrete units or resources rather than as complex geographical and ecological totalities. By reducing the elements of the environment to “resources,” their functions and values could be separated from their connections to each other. For example, water is generally considered as a quantified input for use in irrigation and for household consumption, and forests and trees are seen as simply a supply source for timber. Yet
these elements are not simply abstract resources but form critical environments for numerous species. By separating such elements from their context, the connections that bind them together are silenced, and in the process, much is lost.

Despite the assumed separation of humans from the environment as managers, dominant discourses typically see the elements of the environment only in relation to themselves, as if they existed independently of human intervention and modification. For example, in Australia, legislative definitions of the environment never allow the biophysical environment to be defined in isolation; it is always defined in relation to people, and therefore, the social aspect is always present. While there is a large body of work about biophysical relations and connections, these understandings must move beyond the biophysical realm to consider its elements in light of their relationships to human beings and social structures. Indeed, many elements of the environment are framed as “natural resources” that are valued and understood for their contribution to humans. In the process, environmental management is easily conflated with natural resource management, as if the governance of human relationships with place is best achieved in terms of managing resource values. For example, governments worldwide have prioritized agricultural, forestry, and the soil sciences to focus on the potential of primary industries, and environmental assessments for proposed developments are often done with respect to how they serve dominant political or economic purposes. Implicitly, the inextricable connections between humans and their environs still exist within these approaches, yet dominant discourses of environmental management position the “environment” as separate from people in a manner that facilitates the use, commodification, and often exploitation of the natural world.

Connectivity and Relationships
Universalizing scientific discourses have often silenced other approaches concerning human relations to the environment. For example, many ecologists, indigenous peoples, and anthropologists and geographers, and even some environmental managers, argue for a richer and more robust version of environmental management that goes beyond the simple engineering of abstracted resources toward an understanding of the relationships and connections between species and, critically, the embedded nature of humans within these relationships. Such approaches do not view people and the environment as separate, disconnected worlds but rather as two domains inseparable intertwined with one another.

Indigenous perspectives on environmental management typically provide a clear alternative to dominant discourses by acknowledging the connections between, rather than the separation of, nature and culture, a view in which humans are not considered simply as managers. Such views do not perceive environmental “management” as human actions on the environment. Instead, they articulate such activities as taking place within and being accountable to an intricate web of relationships and processes in which humans have important relationships with their environment and, as such, rights and responsibilities in a form of mutual obligation. Indeed, many indigenous groups recognize the role of environmental factors beyond just human needs to include the rights and agency of other species. Indigenous approaches to understanding management often position human activities within the webs of the relationships in which they are embedded, recognizing the inherent codependence of nature and culture. Humans are not separated from, or positioned as superior to, their environments. Their actions are performed as part of an intersecting group of reciprocal rights and responsibilities. There are increasing examples of the ways environmental management can be reconsidered to explicitly incorporate the importance of human and environment relations.

The Great Whale River Hydro-Electric Project, which was proposed for development in Northern Quebec, Canada, in the early 1990s, provides an example of the development of innovative assessment guidelines that explicitly sought a multicultural definition of the environment. The guidelines acknowledged more than simply the ecological components of the proposed development, to include the ethical and practical aspects
as well. While working to understand the various ecosystem elements, they aspired to something beyond a simple encyclopedic account. The focus was on what was referred to as “valued ecosystem components” and recognized that the relevant values were constructed within and were reflective of a “multicultural definition of environment,” which took seriously the concerns, fears, aspirations, and values of all the affected communities (including but not limited to scientists). This approach required an assessment of the environmental effects of the proposal in terms of important scientific, cultural, and health aspects that needed to be understood, documented, and managed to ensure social and ecological sustainability.

The “cultural flow” approach to water management is another innovative approach that recognizes and respects the multiple interests and relationships within environmental systems. Focusing on cultural flows encourages a consideration much broader than the commercial networks of water or simply the constituent elements of the systems. Relationships between humans and place, the core of geographic analysis, are made explicit in these approaches. These examples provide windows of opportunities for envisaging alternative approaches to managing the environment, focusing on the importance of process, engagement, and relationships rather than the abstraction, objectification, and intervention that Eurocentric science provides for within prescribed outcomes.

Conclusion

Albert Einstein once said that we cannot solve problems by using the same kind of thinking we used when we created them. Eurocentric scientific thinking, in its efforts to make sense of the world, has frequently focused on the ways in which the environment can be understood and managed as separate, independent components of a complex whole. Dominant approaches to environmental management have created a large and important accumulation of information about the biophysical nature of the environment and the impacts of human behavior on it. However, in doing so, humans are discursively positioned as separate from and superior to the environment, which has allowed for the abstraction of elements, such as water, to be quantified, justifying intervention and management control by humans. This separation is in fact what has enabled and encouraged natural resource allocation and exploitation, which at times has led to overallocation and overexploitation. The dominance of Eurocentric science as the appropriate response to environmental degradation has meant that alternative views that focus on connections and relationships have often been ignored. Thinking differently about the ways in which humans relate to their environments, the core of geographic analysis, may provide new ways of addressing today’s environmental problems.

Samantha L. Muller and Richard Howitt

See also Community-Based Conservation; Conservation; Drylands; Eurocentrism; Governmentality and Conservation; Indigenous Environmental Knowledge; Nature; Nature-Society Theory; Resource Economics; Resource Geography; Resource Tenure; Social Construction of Nature; Soil Conservation; Spatial Strategies of Conservation

Further Readings


Dryland environmental management systems comprise the individual and collective strategies used by resource users to preserve the biological productivity and the economic and social benefits derived from dryland environments. Approaches to managing dryland environments are as diverse as the environmental and social factors that shape human use and management of dryland resources. Drylands encompass a gradient of moisture-deficient zones that span the world’s hyperarid, arid, semiarid, and dry subhumid areas. Ecological patterns within drylands exhibit considerable diversity, incorporating woodlands, grasslands, scrublands, and deserts. This diverse ecology supports a range of human livelihood systems, including dryland farming, pastoralism, extractive activities, and large-scale urban settlement. Because dryland resources support the livelihoods of a disproportionate number of people living in extreme poverty as well as growing populations in dryland urban centers, the sustainable management of dryland environments constitutes one of the major global environmental challenges of the 21st century.

Whereas specific management regimes vary by environmental conditions and patterns of resource use, dryland management is typically adapted to extreme temporal and spatial variability of precipitation and water resources and, consequently, variability in patterns of vegetation and wildlife. As a result, human management systems emphasize resilience in the face of climatic extremes and uncertainty. Mobility, flexibility, and opportunism are hallmarks of many local systems of dryland management. Dryland environmental management often involves managing human livelihood activities in relation to limited water resources while addressing the interrelated problems of loss of vegetative cover, soil erosion, salinization, and loss of biological diversity.

Patterns of environmental management reflect the range of institutions, knowledge, practices, and capacities for realizing environmental and social objectives in dryland environments while limiting the impacts of human production systems. Although drylands are often thought of in terms of their biophysical constraints, dryland environments provide a wide range of services that support the increasingly diversified livelihoods of communities within and beyond dryland areas. Dryland environmental management is thus deeply intertwined with the broader livelihood strategies of dryland populations.

Historically, dryland societies have developed sophisticated strategies for coping with arid conditions and extreme variability. Such coping mechanisms enable pastoralist and dryland farming societies to maintain access to key resources, such as groundwater and surface water for livestock, irrigated agriculture, and human consumption (see first photo). Coping mechanisms may also reflect the economic and political pressures placed on dryland societies and the interaction of dryland communities with state and civil society actors.

In the hyperarid and arid realms, pastoralism is the dominant livelihood system. Pastoralist management strategies have historically relied on the mobility of people and livestock within the ecosystem to gain access to water and vegetation for themselves. Such transhuman movements typically involve movement to upland, wetland, or riparian areas during dry seasons and a return to drier plains during the wet season. Movements further afield may be required during droughts. The seasonal shifting of pressure on vegetation resources allows for regeneration of grasslands because overgrazing and loss of vegetative cover can lead to a series of additional problems, including reduced surface water, reduced groundwater recharge, and severe soil erosion. Pastoralists may also modify savanna ecology through biomass burning or tree planting. For example, silvopastoralism can provide alternative sources of fuelwood and charcoal for dryland communities.

Dryland farmers engage in a variety of soil and water conservation strategies to enhance moisture and soil fertility. In the case of rain-fed agriculture, water harvesting is a critical technique for maintaining crop productivity. Farmers have historically relied on local varieties of
drought-resistant crops, such as millet and sorghum in Africa and South Asia. Soil management is equally important as it determines the extent to which moisture is retained to enhance crop productivity. Mulching, stone bunds, agroforestry, and many other techniques are used by dryland farmers around the world to minimize soil erosion. Where surface water or groundwater resources permit, irrigated agriculture limits the impacts of drought on productivity.

Although dryland management systems are attuned to local environmental conditions, patterns of dryland resource use reflect the roles of dryland areas in broader regional political economies. Just as dryland resources are used by populations in humid zones (e.g., mineral, fossil fuel, and woodfuel resources), dryland populations may also rely heavily on external resources such as water. Cities located in the dryland realm, such as Los Angeles, Tucson, and Mexico City, demonstrate the potential for large-scale infrastructure to sustain large populations in dryland areas. However, such urban growth faces many limits in terms of long-term water availability due to political uncertainties and physical limitations.

Institutional Dimensions of Dryland Management

Although cases of widespread dryland degradation have been documented throughout history, desertification rose to the status of a global environmental problem relatively recently as a result of international attention to the prolonged drought in the West African Sahel from 1968 to 1973. International concern culminated in the

Maasai women of Kenya return home with water for domestic use. Local water harvesting and conservation initiatives can buffer communities against the effects of drought.

Source: Author.

At the national level, patterns of dryland environmental management often reflect tensions between local management systems and systems of state and private management, as well as conflict over access to dryland resources. In many cases, dryland areas are economically and politically peripheral or are frontier areas over which states have relatively little political or administrative authority.

Due to the highly variable nature of dryland environments, resource access has historically been subject to tenure systems that provided flexible and negotiated access to common property resources—rather than exclusive rights—within dryland landscapes. Many governments have adopted preservationist policies toward dryland areas as a means of limiting the impacts of human activities on vegetation, soils, and wildlife. The creation of protected areas, such as the extensive national parks systems in Eastern and Southern Africa, has often resulted in the displacement of pastoralist and dryland farming communities and the disruption of traditional management systems. Such policies have often ignored beneficial interactions between livestock, vegetation, and wildlife.

Community-based natural resource management in drylands has been a more recent development and seeks to maintain ecosystem function while recognizing the validity of local knowledge of appropriate environmental management and the complexity of local claims to resources. While environmental monitoring may improve the targeting of environmental policies and programs, broader commitment to community-based resource management may be the most promising means of improving dryland environment management in the face of greater uncertainty associated with global climate change.

Thomas A. Smucker

See also Biome: Desert; Climate: Dry; Common Property Resource Management; Desertification; Drought Risk and Hazard; Indigenous Water Management; Nomadic Herding; Water Needs; Xeriscaping

Further Readings


The process of collecting and visualizing spatially referenced data about our environment is referred to as environmental mapping. Environmental mapping does not refer only to an end product, such as a map of deforestation, but to the whole process of acquiring, preprocessing, storing/archiving, analyzing, and finally displaying information about our physical environment (Figure 1). As such, environmental mapping can be regarded as multidisciplinary. It draws on the expertise of environmental scientists, such as physical geographers, geologists, ecologists, and climatologists, who help define the aspects of the environment to be studied and how to interpret the data collected. It also draws on surveyors, photogrammetrists, and remote sensing specialists to provide expertise in the acquisition of data on the environmental phenomena being studied. Cartographers, modelers, statisticians, graphic designers, and environmental scientists help in the processing and eventual visualization of the environmental data. Environmental mapping can be regarded as underpinning much of the investigation of the physical environment, which is where its importance lies. Scientists and decision makers need to know about the spatial distribution of phenomena in order to be able to research them and
ENVIRONMENTAL MAPPING

develop policies to address environmental problems such as deforestation or climate change. Environmental mapping also has an important role in informing the public and facilitating informed public debate. Environmental mapping is rarely the last stage in the process; it is part of the process of gathering and presenting information for further analysis and interpretation. This stage of investigation then leads to the development of explanations and, when relevant, to the formation of policy and legislation.

The process of environmental mapping consists of a number of stages (Figure 2). The first stage in the process is defining what aspects of the environment are to be mapped. This is then followed by a data collection/acquisition phase. Next, the data may need some form of preprocessing before they can be used. In the case of large projects, or ones where data collection must be repeated numerous times to develop a time sequence of information, some form of data storage or archiving is required. Again, depending on the nature of the data collected and the aspect of the environment being studied, some form of analysis may be required before the data can be displayed for interpretation. The final stage in the environmental mapping process is to create a visual display of the data.

Stage 1: Problem Definition

The first stage in environmental mapping requires making a decision on which aspect of the environment needs to be mapped and what data are required for mapping. This stage will usually be informed by the requirements of the scientific methods behind the project and/or by the policy needs. Ideally, data would be collected for each environmental mapping project, but this is not usually realistic or possible. A key element in the problem definition stage is to assess the fitness of different potential data sets for studying particular environmental phenomena. It is worth noting that the data fitness category is not always limited to identifying the very best data set for a particular application; quite often, it includes criteria such as the data acquisition cost and the availability of resources needed to obtain the data.

Stage 2: Data Acquisition

Once it has been determined what data are to be used, the second stage is data acquisition. This stage of environmental mapping can vary depending on whether existing data sets are to be used or whether new data need to be collected. If existing data sets are to be used, this stage might involve something as simple as downloading data from an Internet site. An example of this might be where a project requires remotely sensed data, which might be accessed from an online archive, such as the one available for satellite remote sensing instruments such as MODIS (moderate resolution imaging spectrometer). The development of the Internet has led to an explosion in the volume of environmental data available, much of it free to download. However, sometimes new data must be collected. In these cases, researchers can use a wide range of technologies and techniques developed specifically for collecting spatially referenced data. Data collection might take the form of direct measurements, such as ecological surveys or measurements made with global positioning systems, or indirect measurements, where remote sensing instruments are used. Depending on the nature of the project, data acquisition might involve a major national or international campaign, observing the whole planet, such as NASA’s Earth Observing System, or it may just
involve an individual going out with a global positioning system to acquire the location of pollution sites or endangered habitats. As indicated in Figure 2, once the data have been collected, there are a number of stages before the final visualization is created.

**Figure 2** The stages in the process of environmental mapping. Some of the stages in the process, such as problem definition (Stage 1), data acquisition (Stage 2), and visualization (Stage 6), are always carried out. Other stages, such as preprocessing (Stage 3), data storage/archiving (Stage 4), and analysis (Stage 5), are carried out when the mapping project requires them.

Source: Author.

making the data usable so that it can be analyzed and/or visualized.

**Stage 4: Data Storage/Archiving**

The storage of the data in an accessible and logical manner is important, particularly if the environmental mapping project being undertaken continues for a period of time. The storage of large quantities of geographically referenced data has become much easier with the advent of geographic information systems (GIS) and related database technologies. The key issue in the storage of data is determining who will be allowed access to the data and what they are allowed to do with it. These are key concerns if a number of people are involved in a project. The Internet has made the sharing of data relatively easy, but protocols are needed to ensure that the data remain safe. There is also a need for an efficient indexing system that allows the wide range of users who may access the data to search for the data sets they require.
Stage 5: Analysis

The preprocessing of data might be sufficient to make data usable, but in some cases additional analyses might be needed to obtain information on the phenomena or problem being mapped. The analysis might involve the combination of a number of data sets, such as combining layers of information on slope angle, rainfall, and vegetation type to create a model of soil loss. The analysis might involve running hydrological models to create different flooding scenarios. The purpose of analysis is to create information that can be visualized. As noted earlier, the environmental mapping process may be a part of a larger investigation, and so the analysis done as part of environmental mapping may be only a starting point to create derived data and information useful in more complex modeling and simulation.

Stage 6: Visualization

Generally, the final step in the process of environmental mapping is the visualization of the data. Originally, visualization was limited to static two-dimensional representations, usually in the form of a thematic map. However, modern technology makes it possible to visualize environmental entities and their relationships in three dimensions or even to animate entity and relationship changes in time, thus providing the means to visualize processes. Visualizations can comprise a wide range of different data types, including traditional maps, digital elevation models, remotely sensed data, aerial photography, and even virtual reality objects. Technologies such as GIS and the Internet have aided the visualization component of environmental mapping greatly. GIS has provided the framework for storing and displaying the data, and the Internet has made it possible to share that information more widely.

Examples of Environmental Mapping

Environmental mapping can be carried out at many scales, from a very local scale right up to a global scale. A number of examples of environmental mapping–based applications can be found on the Internet. In the United Kingdom, the Environment Agency provides an Internet-based service where the general public can look up information on their local environment using postal addresses. Information on flooding, water quality, water management, and pollution can all be looked up and displayed as thematic maps. This service provides a basic level of environmental mapping. A similar type of application can be found on the Web sites of the U.S. Environmental Protection Agency (EPA) and the U.S. Department of Homes and Communities. On EPA’s Web site, the EnviroMapper tool allows the general public to produce interactive maps of various aspects of the environment, including pollution and flooding. Although these examples describe the final stage in the environmental mapping process—visualization—they demonstrate one of the key roles of an environmental mapping project, namely providing a visual record of some aspect of the environment.

Richard Peter Armitage

See also Adaptation to Climate Change; Aerial Imagery: Data; Aerial Imagery: Interpretation; Biophysical Remote Sensing; Cartography, History of; Climate Change; Database Management Systems; Data Query in GIS; Digital Terrain Model; Digitizing; Ecological Mapping; Forest Degradation; Geostatistics; GIS in Environmental Management; GIS in Land Use Management; GIS Software; Internet GIS; Map Visualization; Remote Sensing; Resource Mapping; Surveying; Three-Dimensional Data Models

Further Readings

Originally developed by geographers and psychologists in the 1950s through the 1970s, environmental perception focuses on how people sense, mentally process, and act on patterns they perceive in space and time. In truth, there is no simple definition of environmental perception in that the two components of the concept—environment and perception—are anything but simple. The environment can be defined as the reality around us in which people live and act. William Kirk introduced into geography a dual view of the environment from gestalt psychology of the phenomenal environment, which lies outside a person’s perceptions, and the behavioral environment, which is the perceptual environment, in which the phenomenal environment is given meaning through the senses, experiences, culture, and values of the individual. David Lowenthal introduced environmental perception as the medium by which people perceive the reality that they experience and within which they act. He divided the universe of geographic study into three realms: (1) the nature of the environment, (2) what we think and feel about the environment, and (3) how we behave in the environment. The Earth science tradition in geography has focused on developing a strong understanding of the nature or “reality” of the environment. Environmental perception is focused on Lowenthal’s second and third realms.

Perceiving the Environment

Humans perceive the environment through their senses. The senses by which we perceive the environment include the hepatic (touch), olfactory (smell), auditory (hearing), and visual (sight). Senses vary in terms of the range of space that can be perceived: The hepatic zone is the immediate perception of differences of pressure and temperature on our skin, while the olfactory, auditory, and visual zones vary with the strength of the surrounding stimuli and the acuity of individual senses. Each person perceives the environment within a personally developed “filter” that incorporates personal history and experiences, which may also vary with mood, purpose, and focus. The “filter” through which we perceive our environment is shaped through culture, custom, beliefs, and desires.

Reginald Golledge and Robert Stimson outlined the problems in defining perception and how it relates to cognition. Some geographers use perception to describe how aspects of the environment are remembered or recalled. In planning and architecture, perception can refer to common interests between groups and individuals in the design process. In psychology, perception is viewed as a subprocess of cognition. Perception is the immediate stimulus-dependent interpretation of the environment through the senses, while cognition refers to how we link that initial interpretation with past experience and how we may project that interpretation into the future. Robert Lloyd defined perception as the cognitive process that is directly involved with the detection and interpretation of sensory information. Regardless of the varying ways in which researchers in different fields define perception and differentiate between perception and cognition, it can be agreed that perception involves how the mind interfaces with the environment.

Perception also means understanding. Environmental perception under this definition extends the scope of the concept from the act of sensing the environment to the understanding of the environment gained by integrating that which we sense. It also extends into the decisions and actions that people take based on that understanding and thus blends into the study of spatial behavior or behavioral geography. Environmental perception has developed as a field of geography incorporating both the perceptual and the cognitive processes in human relationships with the environment. As a result of the multiple definitions of perception outlined above, many geographers use perception and cognition interchangeably, along with the terms environmental perception and spatial cognition.

Researching Environmental Perception

In defining environmental perception in geography, one is tempted to apply former U.S. Supreme Court Justice Stewart Potter’s famous statement about pornography in *Jacobellis v. Ohio* (1964):
"I can’t define it, but I know it when I see it.” We can also extend that definition to “I know who does it.” The Environmental Perception and Behavioral Geography (EPBG) Group of the Association of American Geographers has been discussing and researching environmental perception since its inception. The mission of the EPBG is to advance the theoretical and applied interests of environmental perception and behavioral geography within the discipline of geography, developing links to related disciplines through communication and organization. EPBG is a broad subarea within human geography that takes a disaggregated approach to the study of human activity, culture, and society. It is concerned with a diverse set of issues about human behavior, perception, attitudes, beliefs, memory, language, intentions, reasoning, and problem solving involving space and place. EPBG researchers employ a wide array of research methodologies, both qualitative and quantitative, and have interdisciplinary contact with psychology, anthropology, history, phenomenology, microeconomics, computer science, literature, and other disciplines.

Another way of looking at environmental perception in geography is to consider where the research is published. Papers on environmental perception are well represented in mainstream geography journals. Research is also published in the more specialized geographic information science and cartographic journals. In addition, researchers in environmental perception have been publishing in journals in related fields such as psychology, hazards, and planning and architecture, among others. The wide range of publication outlets in these disciplines reflects the multidisciplinary flavor of environmental perception research as well as the multiple perspectives (and definitions) of what research in this field entails.

There are several basic problems that researchers and thinkers in geography and related fields have been investigating and exploring in the area of environmental perception. One of them is how humans perceive the environment. As stated above, a combination of senses and sociocultural experiential filtering is used to perceive and interpret the environment in which we live and act. One variable that has been of particular interest is perspective. Each individual perceives the environment differently, sensing, processing, evaluating, storing, and using information received about the environment. Researchers have been investigating diverse perspectives to determine the differences in perception. Some of the perspectives under investigation include age (both children and the elderly), gender, and disabilities.

Another major area of investigation is the difference between perceptions and “reality.” Since individuals have different perceptions of the environment, geographers have tried to investigate these differences and determine some of the mechanisms behind them. The study of mental maps provides insight into how experience changes our understanding of the environment over time. They also give us some insight on how we visualize space. The perception of space varies with the scale of the space perceived. The results of investigations into the relationship between spatial visualization and environmental perception are being applied to spatial learning through cartography, imagery, and geographic information systems.

Our perceptions of the environment also create feelings for places and an aesthetic value attached to those places. This feeling for specific places as a result of our perception and use of these spaces is referred to as sense of place. Yi-Fu Tuan wrote of the combination of perception, experience, and cultural attitude, which he defined as “topophilia”—the affective bond between people and a place or setting. Architects, landscape architects, planners, psychologists, and geographers have researched the sense of place, the design of places, and the influence of culture, gender, and age on the perception of and feelings for places. Research has also been done on how we value our environment and how our perceptions of the environment physically and mentally restore us.
Environmental perception is also at the heart of the investigation of spatial decision making. How we perceive the environment affects how we act within it. For example, transportation researchers have a strong interest in navigation. How we orient ourselves through space is directly related to how well we perceive the spaces through which we navigate. The cumulative perception and cognition of the environment and representations of the environment create a mental map by which we relate, orient, and navigate spaces. Path selection in space is based on our perception of the “best path,” which can be a function of aesthetic perceptions as well as economic utility. The chosen path is not always the shortest path but could also be the path that is the most aesthetically pleasing. Researchers have examined the impacts of culture, gender, age, disabilities, experience, and environment on navigational abilities and orientation.

How we perceive the environment directly relates to the decisions we make about our environment and the actions we take within it. Neoclassical economists typically assume that humans make rational decisions based on their self-interest; however, humans often make what seem to be irrational decisions. One reason for these seemingly irrational decisions is that we tend to base our decisions on locally scaled perceptions of the environment (both in space and time), where perceptions of the environment differ with scale. One example of this tendency is the behavior of humans with regard to climate change. Since the most obvious impacts of the early stages of global warming are evident only in the polar or alpine regions, the majority of people who live in temperate or tropical climates often do not see significant evidence of climatic change in their local landscapes, and they do not notice significant changes locally from year to year. People therefore tend to discount the impacts of climate change on the basis of local perceptions, and they postpone changes in individual behavior (e.g., reducing individual carbon loads through energy efficiency) that could collectively have positive global impacts.

This contrasts with an alternative case where human behavior was modified through a campaign focused on local environmental perceptions. Littering, like climate change, is a global problem, but littering can be easily perceived (and policed) at the local scale. The Keep America Beautiful campaign of the 1970s, with the iconic image of the “tearful Native American,” expanded individual perceptions of litter from a local to national scale in such a way that individual local behavior was changed nationwide. In each of these cases, people are behaving seemingly rationally based on individual local-scale environmental perceptions, although in the former case they are acting in a seemingly irrational manner in relation to larger-scale, long-term impacts that are not directly perceived.

Other examples of this relationship between local and large-scale perceptions that geographers have explored in depth are decisions involving natural hazards (fire, flood, landslide, earthquakes, volcanoes, hurricanes, etc.). People frequently move to, or even move back to, hazardous locations in spite of the risks involved. The actors perceive the benefits of such locations and downplay the risks, especially as time passes after the most recent disaster. Geographers have also been researching the perception of risk and fear in the landscape. We react to perceptions of natural hazards as well as human and societal hazards, such as crimes and violence. Again, people often react to their perceptions of risk. Those perceptions may vary significantly from the actual risks across the landscape.

The future of environmental perception in geography is bright as interdisciplinary research in the field progresses with new tools, new insights, and continuing cross-fertilization with other fields.

Dave Lemberg

See also Behavioral Geography; Childhood Spatial and Environmental Learning; Environmental Imaginaries; Golledge, Reginald; Humanistic Geography; Mental Maps; Natural Hazards and Risk Analysis; Phenomenology; Sense of Place; Spatial Cognition; Spatial Cognitive Engineering; Symbolism and Place; Tuan, Yi-Fu; Wayfinding
Further Readings


ENVIRONMENTAL PLANNING

Urban planning is often referred to as the ordering and design of human settlements. Accordingly, environmental planning then becomes the consideration and analysis of how existing and proposed human settlements affect ecosystems, with subsequent assessment of the extent to which these impacts can be avoided and/or mitigated now or in the future. As our awareness of human impacts on the environment continues to expand, the practice of environmental planning will continue to take on new importance because it provides an important conduit for scientific knowledge to inform policy development and ultimately potentially change human decision making and behavior, leading to better environmental outcomes. While environmental planning is complementary to the study of environmental geography, it differs from the latter and efforts to respond to sustainable development imperatives. These differences evolve, in part, because of the role that professional urban planners play in the practice of environmental planning.

The History of Environmental Planning

The practice of environmental planning has been influenced by the environmental philosophy of the day. John Randolph has identified three widely recognized phases of evolving environmental values that have fundamentally influenced the practice of environmental planning: the conservation movement, public health, and the ecosystem approach.

The Conservation Movement

The later half of the 20th century gave rise to a philosophical interest in “conserving nature.” Informed by the earlier work of Henry David Thoreau, Aldo Leopold, and John Muir and implemented by nongovernmental organizations such as the Sierra Club, governments began to consider what legislative tools could be employed to protect and preserve vast tracts of undeveloped land or wilderness and the natural resources they held. During this period, environmental planning concerned itself with land outside urban settlements and was applied to vast tracts of land. These preserved or protected areas stood in stark contrast to the increasingly dense urban areas that were home to the infrastructure of the Industrial Revolution.

The Public Health Movement

Responding in part to the negative human health impacts resulting from widespread industrialization located in urban areas, environmental planning’s focus next turned to matters of public health. This movement forged a link in lawmakers’ and decision makers’ minds about the connections between human health and the surrounding environmental conditions. During this phase of environmental planning, there was new urban investment in infrastructure in support of public health, facilitated in part by expanded technological capacity. For example, there were major urban investments in water and sewage treatment, and cities began to explore opportunities to bring “nature” back through the provision of public parks and open spaces.

The transition from this movement to the ecosystem-approach movement was triggered in part by two events. The first was when humankind was first able to view Earth from space—an event widely credited with contributing to our understanding of the planet as a series of connected systems. The second was the release of Rachel
Carson’s seminal book *Silent Spring* in 1962, which sensitized us to the impacts that human activities have on ecosystems.

### The Ecosystem Approach

Planning based on ecosystem approaches began in the early 1980s in North America but gained prominence in other regions after the United Nations Conference on Environment and Development (the so-called Earth Summit) was held in Rio de Janeiro, Brazil, in 1992. Rather than focusing on the environment as a source of resources or a sink for waste, the ecosystem approach to land use planning emphasized decision-making processes that integrated the consideration of the impacts of human development on land, water, and other resources. Through efforts to assess the impact of current development on future generations, the ecosystem approach requires urban planners to think about the impacts of resource use (or conservation) on current and future generations.

---

**Environmental Planning Is Distinct From Environmental Management**

Environmental planning differs from environmental management in part because of the role that the professional urban planner plays. Urban or town planners in many countries are regulated by a professional accreditation body (e.g., The Royal Town Planning Institute, United Kingdom; The Canadian Institute of Planners; The American Institute of Certified Planners, United States; The Institute of Town Planners, India; Planning Institute of Australia). This accreditation process is relevant to the field of environmental planning because it requires its member-planners to adhere to a code of ethics or practice. In many cases (Canada, United States, United Kingdom) this code includes a requirement to uphold ecosystem principles in the practice of planning. For example, the American Institute of Certified Planners’ Code of Ethics and Professional Conduct states that planners “shall promote excellence of design and endeavor to conserve and preserve the integrity and heritage of the natural and built environment.” By meeting this requirement, the planner can influence the practice of environmental planning through content and process. Many nonplanners are engaged in research and practice that generate content for environmental planning processes, but their responsibility for the planning process distinguishes planners from other allied professions.

The process by which environmental planning takes place is typically defined by enabling legislation across national, state/provincial, and local governments. The extent to which the local government holds significant power in the environmental planning process varies across countries. Typically there are multiple jurisdictions involved in the regulation of development activities that have an impact on the environment. For example, for a new subdivision to be approved, the developer may be required to complete an environmental assessment (typically mandated by federal or state/provincial governments), complete a plan of subdivision submitted to the local government, and demonstrate compliance with growth management policies (typically mandated by federal or state/provincial governments).

The environmental planning process also typically must include some form of public consultation whereby the public’s views on a development proposal are solicited by a planner working for government through a public meeting, submission of a letter, or other means of soliciting input. This public consultation process provides formal opportunities for citizens’ groups to raise matters of concern and suggest alternative approaches. Increasingly, there is a growth in the use of online consultation tools to expand opportunities for public comment. Also, citizen groups are turning more to Internet-based spatial analysis tools such as Google Earth and Google maps to create their own sources of data to provide input into the environmental planning process. The increase in the use of public participatory GIS (geographic information systems) tools is likely to continue to grow as these tools proliferate on the Web, the ease of use expands, and the opportunities to integrate multiple data sets expand as well (e.g., mashups).

---

**Current Issues in Environmental Planning**

Although the roots of environmental planning were in the protection of undeveloped areas,
environmental planning is increasingly responding to issues in urban, suburban, and exurban environments. As the practice of environmental planning has internalized the ecosystem approach, the challenges arising from attempting to manage suburban growth have become more apparent. Thus, much attention is paid in environmental planning to the protection of agricultural land from suburban sprawl; reducing greenhouse gas emissions resulting from automobile use; improving local air quality, also degraded by air pollution from automobile use; protecting water quality from surface runoff contamination (exacerbated by the proliferation of parking lots and roadways); and protecting biodiversity and preventing habitat fragmentation and destruction, again as a result of low-density, single-land use urban sprawl.

“Smart growth” is the catchphrase of an approach to environmental planning that seeks to respond to the aforementioned issues. Although different jurisdictions may vary in their interpretation of the concept, smart growth principles generally include mixing recreational, residential, retail, and commercial land uses; building new developments that are compact enough to support public transit and other non-automobile-based travel; providing a range of housing types (e.g., single-family homes, town houses, apartments, or condominiums); redirecting growth to underdeveloped areas (as opposed to undeveloped or greenfield lands); protection of open spaces and environmentally sensitive land; protection and enhancement of agricultural land; using “green” infrastructure and green building practices; encouraging the uniqueness of places; and providing meaningful opportunities to engage the public in the planning process. A review of these principles reveals that there are multiple environmental benefits to be gained through the practice of smart growth-type environmental planning. Accordingly, this concept permeates North American planning policy. New urbanism and transit-oriented design are two similar yet distinct planning frameworks that operationalize the principles of smart growth with different emphasis.

In addition to smart growth, other current trends in environmental planning involve exploring how communities can reduce their ecological footprint (the amount of land required to sustain a person’s lifestyle), reduce greenhouse gas emissions, increasingly localize their food production process, and engage in the revitalization of old industrial sites (brownfields). These issues are all indicative of the elements of environmental planning that distinguish its practice from environmental management or research conducted by environmental geographers. Urban planners who practice environmental planning make significant use of research developed by environmental managers and geographers to inform decision making in the planning process. But planners are also required to mediate interests, negotiate between disparate stakeholders, act as liaisons with politicians, and develop and implement public consultation activities. Environmental planning draws heavily from the spatial analysis that results from the use of remote sensing and GIS, but these data are inputs into a political process that results in mediation between often competing interests.

One persistent and unresolved question in environmental planning arises when the concept is compared with that of urban sustainability. Research and practice related to urban sustainability emerge as a subset of the broader issue of sustainable development. With sustainable development’s emphasis on considering the social, economic, and environmental impacts of human developments for current and future generations, there are obvious synergies with the aforementioned ecosystem approach. As a result of these similarities, the terms urban sustainability, environmental sustainability, and environmental planning are often used interchangeably in research and practice. The extent to which environmental planning can be considered the same as planning for urban sustainability depends in large part on the mind-set of the government or practitioner using the terms. In some contexts, environmental planning focuses on ecological impacts (e.g., loss of biodiversity in wetlands), and the planning outcomes respond to these alone (e.g., new zoning requirements for on-site storm water management to reduce overland flow that compromises wetland water quality). In other
cases, environmental planning takes a broader approach that explores the dynamics between social, economic, and environmental impacts and intervenes accordingly (e.g., the creation of buffer zones between new developments and ecologically sensitive areas to reduce the impact of such developments on wetlands).

While the practice of using these terms interchangeably is widespread, research suggests that there are costs to the indiscriminate use of these terms. The most common critique of this practice emerges from a growing body of research that suggests that the equity element of sustainable development gets lost in the translation from definition to practice, with the environmental and economic elements taking precedence. Also common is the concern that the weaknesses of the concept of sustainable development will be used as an argument against all planning intervention, thus serving as a barrier to even the environmental elements being addressed.

As public and professional awareness of the environmental impacts of human development expand and as the world becomes increasingly urbanized, an increase in the demand for environmental planning research and practice can be anticipated.

Pamela Robinson

See also Brownfields; Community-Based Environmental Planning; Environmental Ethics; Environmental Impact Assessment; Environmental Impact Statement; Environmental Law; Environmental Management; Environmental Protection; Environmental Restoration; Environmental Services; GIS in Environmental Management; Land Use Planning; Multistakeholder Participation; Regional Environmental Planning; Smart Growth; United Nations Conference on Environment and Development; Urban and Regional Planning; Urban Environmental Studies; Urban Sustainability

Further Readings

contributes to cloud formation, whereas carbon contributes to plant growth. Plants are eaten by animals, which are themselves often eaten by others. Animal wastes become part of the lithosphere through decomposition by bacteria and other creatures to become part of the soil, enriching it with nutrients. In the long run, the tendency is for these processes to interact in a balanced fashion. If there is a disturbance, such as an increase in the population of a particular species, then this system can be thrown off balance as that species depletes its food and water resources and some of its members die until balance is restored.

Humans, like other life forms, play their part in transforming their surroundings and create by-products in the process of procuring their livelihoods. They mine and extract resources, transforming them into goods for consumption. In the process of doing so, they create wastes. As the human habitat expands, it alters land cover and displaces whatever ecosystem was there before. As the human population has expanded, increased its affluence, and developed new technologies, the amount of goods and services that it takes from the environment has increased.

An environmental protection specialist considers the following questions: At what point does the satisfaction of human needs come at odds with what the environment can provide? When this happens, what should be done about it? Can environmental goods and services be used sustainably? To a degree, the amount of protection that can be afforded to the environment has to be justified in terms of costs and benefits. The social benefit of the protection measure has to outweigh the costs it imposes on the polluters. While this approach is extremely helpful in environmental protection, one problem with it is that while damage to buildings and crops can be valued with the price of replacement, it is very difficult to put a price on an ecosystem’s goods and services. Ultimately, all areas that are affected by human-caused damage affect the global environmental balance either directly or indirectly through ripple effects throughout several ecosystems and spheres. However, the benefits or costs cannot be measured directly if the environmental damage does not have a direct impact on human health or ecosystems; that is, if the environmental damage has an indirect effect, it is difficult to measure because its effect cannot be readily isolated from other influences and measured accordingly.

The scale of the environmental problem matters for environmental protection. Some environmental protection takes place at a local level. One example is requiring that a pulp and paper facility limit its toxic discharges into nearby water streams. Other environmental protection objectives need a concerted global effort, such as limiting the ozone layer–depleting production of chlorofluorocarbons (CFCs) in many countries via the Montreal Protocol. This is because, as in the case of CFCs, for example, the emissions generated in one country affect the atmosphere, which is shared by the rest of the planet. There is a similar situation with the emission of greenhouse gases; while a few countries may be responsible for a large share of total emissions, the burden of climate change is shared worldwide. Therefore, there can be conflict and trade-offs in environmental protection at different scales.

Many countries have organizations devoted to environmental research and protection. In the United States, the Environmental Protection Agency (EPA) is tasked with implementing environmental law written by Congress. Such laws include the Clean Air Act, which aims to limit the amount of six pollutants released into the air. The EPA sets its National Ambient Air Quality Standards in two ways. A primary standard is designed to protect public health, including sensitive populations (e.g., asthmatics, the elderly, children). Secondary standards are set to protect public welfare (e.g., decreased visibility due to air pollution, damage to crops and buildings). Similarly, for drinking water, EPA has established the National Primary and Secondary Drinking Water Regulations, restricting pollution in public water systems. The U.S. Fish and Wildlife Service and National Oceanic and Atmospheric Administration are charged with the administration of the Endangered Species Act, which is designed to protect different species from overexploitation that can lead to extinction. There are also national and worldwide nonprofit organizations that publish information on environmental concerns and propose or take actions toward their resolution, such as the World Wide Fund for Nature, Greenpeace, Sierra Club, and National Wildlife Federation.
ENVIRONMENTAL RACISM

Environmental racism refers to the process and outcome of racial discrimination in the enactment and enforcement of environmental laws, rules, and regulations and in targeting minority communities for the siting of polluting industries. It also refers to the exclusion or restriction of people of color from representation in public and private boards and commissions and from meaningful participation in environmental movement organizations. The broader term, environmental justice, refers to environmental inequities in relation to a larger set of social categories, such as social class and gender. Environmental justice also refers to the movement to address environmental injustice and environmental racism. The term environmental racism was coined by Reverend Dr. Benjamin Chavis Jr. of the United Church of Christ (UCC) Commission for Racial Justice. Grassroots environmental justice activists have adopted the term to articulate grievances and as a frame to mobilize supporters and claim the right to a clean and healthy environment “where we live, work and play.”

In 1990, the Southwest Organizing Project and the Gulf Coast Tenants Union publicly charged mainstream environmental organizations with environmental racism, that is, with excluding people of color from their boards, staffs, and memberships and failing to address their concerns programmatically. The growing national presence of the environmental justice movement was manifested at the First People of Color Leadership Summit in 1991, where 17 environmental justice principles were adopted.

Since then, the understanding of racial discrimination in environmental policy making has improved. Environmental racism existed long before it was defined as such. Over time, however, the more blatant forms of discrimination associated with environmental inequality have given way to subtler, institutionalized forms. Recent calls for equal protection and nondiscrimination in environmental decision making have invoked the U.S. Constitution and the Civil Rights Act of 1964 and advanced ideals of environmental rights.

Some theories of racial discrimination posit that minority communities are intentionally targeted for prejudicial reasons, beliefs in racial inferiority, or a desire to protect racial group privilege. The sense of racial entitlement provides motivation, and socioeconomic power provides the means for systematic oppression based on race/ethnicity, which in turn can result in environmental inequalities in the distribution of environmental burdens and access to natural resources and environmental amenities such as clean air, clean water, and parks.

The traditional perspective views discrimination as more individualistic, sporadic, and episodic than the institutional perspective, which accentuates the routine, continual character of discrimination. Racial discrimination also can take an institutionalized form not necessarily related to overt racism when, for example, siting and cleanup decision rules, procedures, or the practices of industry and government lead to racially disparate outcomes, even though they appear race-neutral. For instance, decision-making bodies and their governing rules or the way they are carried out may limit access to information or participation of people of color in hazardous waste facility siting, permitting, and decisions.

Discrimination in various institutional domains such as housing, education, employment, and health care can further disadvantage minorities and limit their social and physical mobility. Such
“structured disadvantage” contributes to inter-
generational poverty and racially segregated resi-
dential land use patterns, which in turn increases
minority sociopolitical and economic disadvan-
tages and vulnerability to targeting.

Commonly cited examples of institutional dis-
crimination include government policies and pri-
ivate investment to support suburbanization,
coupled with disinvestment in inner cities. Since
the 1950s, these policies and practices, coupled
with “white flight,” have led to economic decline
and a decaying infrastructure in the urban core.
Discriminatory real estate and lending practices
and expulsive zoning have further concentrated
racial-ethnic minorities in areas where locally
unwanted land uses (waste facilities in particular)
have been increasingly located. Institutional dis-
crimination in environmental policy allows the
licensing of polluting industrial facilities and the
release of toxic chemicals into the environment
without consideration of the disparate impacts on
minority and low-income populations.

Empirical Evidence

In the 1980s and 1990s, studies testing various
racial explanations for environmental inequality
produced mixed results. Among the most influen-
tial early studies drawing attention to the prob-
lem were a 1983 U.S. General Accounting Office
(GAO) study, the UCC’s 1987 landmark study,
Toxic Wastes and Race in the United States,
and Robert Bullard’s Dumping in Dixie. Although
these studies reported evidence of racial dispari-
ties in the location of environmentally hazardous
sites, other early studies found only small or insig-
nificant disparities, if any. In some cases, racial
disparities appeared to be explained entirely by
differences in socioeconomic status.

More recently, studies have examined causal
ordering. Some studies reported racial disparities
existing at the time of siting; others found evi-
dence of racial disparities developing because of
demographic changes occurring after siting, while
a few found evidence supporting neither process.

Extensive reviews of the rapidly expanding
empirical literature, however, have shown race to
be a stronger factor than income or socioeco-
nomic status in explaining the locations of pollut-
ing industrial facilities. Improved methods for
assessing locational disparities and identifying
affected populations by using geographic infor-
mation systems have shown that racial dispari-
ties are greater than previously found. Using these
newer methods, which more reliably count people
living in close proximity to the nation’s hazard-
ous waste facilities, the report by Robert Bullard
and colleagues, Toxic Wastes and Race at Twenty:
Grassroots Struggles to Dismantle Environmen-
tal Racism, found that race maps closely with the
geography of pollution and concluded that racial
factors are stronger predictors of facility loca-
tions than is the socioeconomic status of host
neighborhoods. These findings are consistent with
a 2005 Associated Press study showing that Afri-
can Americans were more than twice as likely as
white Americans to live in neighborhoods where
air pollution poses the greatest health danger.
Hispanics and Asians were also more likely to
breathe hazardous air pollutants in some regions
of the United States.

Exposure to the heavy metal lead is one type of
hazard known to be particularly high in African
American children, due to their being concen-
trated in older housing with lead paint. Harmful
exposure to mercury and other bio-accumulating
toxins in fish and marine mammals is another
risk that is disproportionately high among Afri-
can Americans, Asian immigrants, and some
Native Americans who hunt and fish for subsis-
tence or to supplement protein in their diets. For
some risks, the combination of being a member
of a minority community and having a low socio-
economic status confers added risks beyond those
associated with being a minority member alone.

The “Smoking Gun”

Following the polychlorinated biphenyls landfill
protests in Warren County, North Carolina, in
1983, the GAO released a report that found that
three of four hazardous waste landfills in the
southeastern United States were located in pre-
dominantly African American communities.
These developments led to charges of racial tar-
geting. While direct evidence of intentional dis-
crimination is nearly impossible to prove, a 1984
report commissioned by the California Integrated
Waste Management Board is widely cited as a
“smoking gun.” Because of the growing antitoxic
movement in middle- and upper-class white communities, siting new waste facilities had nearly come to a halt in the early 1980s. The report, which became known as the Cerrel Report, sought to identify communities less likely to wage not-in-my-backyard campaigns. In developing the profiles of communities likely to be more accepting of waste facilities, the report did not mention race explicitly, but it employed what many viewed as racial code language and served to confirm the suspicions of many in the environmental justice movement.

Highlighted in the 1983 GAO report, Waste Management, Inc. became identified by environmental justice advocates as a company with a track record of locating landfills in minority communities. In the 1990s, a Waste Management subsidiary operated a hazardous waste landfill in Kettleman City, California, where a proposed expansion has been a frontline environmental justice struggle for more than a decade. Elsewhere, the poor performance of some companies with records of accidents and permit violations harming workers and nearby residents has provided a rallying cry for community groups for stronger enforcement, fines, and penalties.

Perhaps more than any other case, Hurricane Katrina, in 2005, revealed America’s racial divides. As documented in a report by the Russell Sage Foundation, Hurricane Katrina has shown the face of institutional discrimination in disaster preparedness and response. The disaster and the increases in hurricane intensity have also heightened awareness of climate justice, that is, concern about the disproportionate impacts of climate change and climate change mitigation on Arctic peoples and atoll nations in particular.

Environmental Racism Remedies

Environmental justice groups, often with assistance from university researchers, students, public health professionals, and faith-based groups, have made great strides in improving environmental conditions in communities overburdened with pollution. Working in diverse coalitions, often by using local knowledge, environmental justice groups make use of citizens’ right-to-know legislation, citizen science, popular epidemiology, cumulative risk assessment, and legal investigations to marshal evidence of disparate impacts. As a result, the environmental justice movement has contributed to pollution prevention, toxic substance use reduction, and modest shifts to cleaner production technologies.

Despite these gains, many of the efforts to formulate policies and use existing laws to remedy environmental racism have met with limited success. In 1998, the U.S. Environmental Protection Agency (EPA) issued draft guidance for state environmental agencies investigating civil rights complaints. Although the guidance was revised in 2000, EPA has never promulgated regulations, and EPA’s Office of Civil Rights has been slow to investigate and resolve civil rights complaints. State governments have also proposed or adopted environmental justice policies, for example, to conduct disparate impact analyses as part of siting and permitting processes so that existing environmental burdens and health status are taken into account.

A few executive orders have been issued to address environmental justice concerns, most notably Executive Order (E.O.) 12898 of 1994, issued by then U.S. President Bill Clinton: “Federal Actions to Address Environmental Justice in Minority Populations and Low-Income Populations,” which required federal agencies to take into account potential disproportionate impacts of their policies and programs on minority and low-income populations. Various strategic plans were developed by federal agencies. However, recently EPA has removed race as a consideration in its Environmental Justice Strategic Plan, and its Inspector General issued two reports critical of EPA’s implementation of E.O. 12898. This has led to recent calls for codification of the order through statute and “environmental communities of concern” to be designated for targeted enforcement and other assistance.

The courts have restricted the ability of disproportionately affected communities to use Title VI of the Civil Rights Act of 1964, which forbids federal agencies from discrimination by “race, color, or national origin.” In its Alexander v. Sandoval decision in 2001, the U.S. Supreme Court denied a private right of action for racial discrimination cases brought under Title VI, thereby requiring the government to carry cases for aggrieved parties. The Court also held that there must be proof that discrimination was intentional,
not just that a federal policy or action had a discriminatory impact based on race or ethnicity. Although toxic tort claims do not seek to enforce civil rights, they are currently being used as an alternate legal remedy.

Assertion of international human rights and indigenous rights is an avenue being used to combat environmental racism using the international judicial system. For example, claims of human rights violations to the Inuit people resulting from global warming have been heard by the Inter-American Commission on Human Rights.

International agreements such as the Basel Convention on the Control of Transboundary Movements of Hazardous Wastes and Their Disposal hold promise for addressing North-South manifestations of environmental racism at the global scale. Nevertheless, illegal “e-waste” and other hazardous waste exports from developed to less developed and developing nations continue due to lax enforcement, and various banned substances also continue to be exported abroad.

Mainstream environmental organizations have begun to respond to these problems. Among the large environmental groups, Greenpeace has vigorously supported environmental justice struggles, most notably under the leadership of the late Damu Smith, with regard to the Shintech Corp. case in Convent, Louisiana, which involved the siting of a large polyvinyl chloride factory. Prominent mainstream groups have made limited progress in diversifying their boards, staffs, and memberships. Nevertheless, several U.S.-based environmental groups, such as Sierra Club, Natural Resources Defense Council, and National Wildlife Federation, have hired staff to work on environmental justice issues. Foundation support for grassroots environmental justice groups continues to lag, with mainstream groups obtaining the bulk of private foundation and corporate grants.

In the United States, environmental justice organizations continue to take matters into their own hands by forming networks and coalitions, such as the Southwest Network for Economic and Environmental Justice, the National Black Environmental Justice Network, and the Indigenous Environmental Network. Although there have been meetings to forge unified multiracial, multi-ethnic, and cross-class alliances, the lack of a national political organization for environmental justice or a lobbying presence in Washington, D.C., may limit the effectiveness of the movement in the foreseeable future.

The apparent intractability of environmental racism has led to calls to fundamentally restructure systems of production. Various quarters have recommended that government and industry adopt policies informed by the “Precautionary Principle,” which reverses the burden of proof of harm from adversely affected communities to industry by requiring it to demonstrate safety before releasing chemicals into the environment or products into the marketplace.

Environmental racism could be reduced by the phaseout of persistent, bioaccumulative toxic chemicals and the adoption of clean production principles and methods with zero-waste and closed-loop systems that use renewable energy, nontoxic materials, safer chemical practices, and sustainable product design. Also offering promise are green procurement policies and Extended Producer Responsibility, which requires producers to take responsibility for the entire product life cycle, including the postconsumer phase.

Robin Saha

See also Brownfields; Ecological Justice; Environmental Justice; Ethnicity and Nature; Hurricane Katrina; International Environmental Movements; Justice, Geography of; Locally Unwanted Land Uses (LULUs); Not in My Backyard (NIMBY); Polychlorinated Biphenyls (PCBs); Race and Nature; Race and Racism; Segregation and Geography; Social Justice

Further Readings


Environmental refugee is a term that has gained currency in civil society and the media to describe people who undergo forced migration linked to environmental change. Although migration linked to deteriorating environmental conditions is not a new phenomenon, the concept of environmental refugees has emerged since the 1970s in parallel with environmental crises, particularly desertification in Africa and climate change impacts in the wider developing world. It is frequently used to describe and highlight how populations that have been displaced or are at risk of displacement associated with environmental change fall outside the ambit of protection provided to those legally designated as refugees by the United Nations High Commissioner for Refugees. The term has been debated and critiqued in academic inquiry rather than accepted as a useful theoretical concept.

**Numbers of Environmental Refugees**

Environmental change is a broad concept, encompassing short-term disasters such as cyclones, tsunamis, and eruption of volcanoes, as well as progressive phenomena such as land degradation, deforestation, and sea-level rise. It is thus difficult to calculate the numbers of environmental refugees. Norman Myers, however, calculated that in 1995 there were 25 million environmental refugees worldwide, 5 million of whom were in the African Sahel. The same study predicted that some 200 million worldwide would be displaced in the future by climate change impacts. Such estimates are frequently questioned on the basis of a lack of policy or scholarly agreement on what environmental refugee status entails.

**Legal Status**

Environmental refugees differ legally from political refugees. The latter are defined in the United Nations 1951 Convention Relating to the Status of Refugees and the 1967 Protocol to the Convention. Refugees are defined there as people outside their state of nationality or former residence who, owing to a well-founded fear of being persecuted for reasons of race, religion, nationality, or membership of a particular social group or political opinion, are unwilling or unable to return to it. Environmental factors contributing to migration fall outside this definition. Hence, there is no obligation for state signatories to the above-named convention to recognize or protect even the most destitute displaced persons in the absence of persecution. Nevertheless, the legal definition of a refugee does not depend primarily on whether or how migration is forced but on the crossing of an international boundary and the consequent need for protection that is not provided by the country of origin. Thus, in circumstances where a
person satisfies the legal criteria for refugee status, the question of environmental degradation as a causal factor is redundant.

**Perspectives on Migration and Environmental Change**

There are two contrasting perspectives on links between migration and environmental change that are relevant to environmental refugees. Maximalist perspectives posit that environmental degradation is a primary cause of population displacements. Direct causal links between the two are legitimized when the concept of environmental refugees is assumed to possess explanatory power in empirical contexts. For example, Norman Myers defines environmental refugees as people who can no longer gain a secure livelihood in their homelands because of drought, soil erosion, desertification, deforestation, and other environmental problems, together with the associated problems of population pressures and profound poverty.

Minimalist perspectives, in contrast, see environmental change as a contextual variable that can contribute to migration rather than being its primary cause. Such perspectives position environmental change in a way that cannot be analytically and empirically isolated from the economic, cultural, and political fabric associated with migration events. Minimalist views take into account the fact that migration can reflect both long-standing customs and new opportunities and challenges in a globalized world. These views draw on political ecology to problematize the links between environmental change and power relations in migration events. Minimalist views are particularly prominent in climate change debate. “Climate refugees” are understood as a specific class of environmental refugee likely to be affected by climate change impacts such as sea-level rise, drought, flood, and resource scarcity. Migration of people away from low-lying island regions, for example, is at times assumed in environmental campaigns and the popular media to be an involuntary flight, long before these regions have become uninhabitable due to sea-level rise.

The popular currency of environmental refugee discourse has prompted some migration scholars to question the interests that the often well-intentioned deployment of the term *environmental refugee* serves, using minimalist perspectives on the links between environmental change and migration. Maximalist perspectives within civil society are critiqued for promoting the interests of environmental refugees as a proxy for addressing the “real” problem of environmental degradation and potentially failing to advance the interests of populations in environmentally vulnerable locations on whose behalf they claim an advocacy role. From a minimalist perspective, the links between migration and climate change focus on culturally and environmentally appropriate adaptation measures, including but not limited to migration options. For the populations of many small islands in the Pacific, for example, migration is a process embedded in cultures in which seafaring has long been central. Employment on commercial ships and multidirectional migration between the islands and Pacific Rim cities (particularly in New Zealand) is common. It is within these contexts of customary seafaring and high levels of mobility that the issue of the loss of habitability of the Pacific islands is being negotiated.

**Politics of Environmental Refugees**

Groups of involuntary migrants crossing international borders can be perceived as negatively affecting the receiving state. Public opinion on perceived impacts on the environment through urban pollution from refugee settlements, on the economy through welfare systems, and on cultural and social norms and practices can conflict with policies on reception and resettlement of displaced people with claims to refugee status. In this highly politicized context, the concept of the environmental refugee is being used in civil society as a form of advocacy for populations who are vulnerable to environmental change, whether or not migration has indeed occurred. The preemptive labeling of populations as environmental refugees is particularly prominent in climate change debate. “Climate refugees” are understood as a specific class of environmental refugee likely to be affected by climate change impacts such as sea-level rise, drought, flood, and resource scarcity. Migration of people away from low-lying island regions, for example, is at times assumed in environmental campaigns and the popular media to be an involuntary flight, long before these regions have become uninhabitable due to sea-level rise.

The popular currency of environmental refugee discourse has prompted some migration scholars to question the interests that the often well-intentioned deployment of the term *environmental refugee* serves, using minimalist perspectives on the links between environmental change and migration. Maximalist perspectives within civil society are critiqued for promoting the interests of environmental refugees as a proxy for addressing the “real” problem of environmental degradation and potentially failing to advance the interests of populations in environmentally vulnerable locations on whose behalf they claim an advocacy role. From a minimalist perspective, the links between migration and climate change focus on culturally and environmentally appropriate adaptation measures, including but not limited to migration options. For the populations of many small islands in the Pacific, for example, migration is a process embedded in cultures in which seafaring has long been central. Employment on commercial ships and multidirectional migration between the islands and Pacific Rim cities (particularly in New Zealand) is common. It is within these contexts of customary seafaring and high levels of mobility that the issue of the loss of habitability of the Pacific islands is being negotiated.
by their inhabitants. For many of these people, the prospect of future environmental refugee status is only one issue among a host of social, economic, and environmental concerns linked to migration.

Carol Farbotko

See also Climate Change; Desertification; Migration; Political Ecology; Refugees

Further Readings


Environmental restoration is an increasingly common management strategy in physical and ecological systems that are viewed as degraded or impaired. Although restoration of environmental systems is defined in different ways by different practitioners, it has grown considerably in recent years in both practice and theory. The term environmental restoration itself also represents a complex series of abstractions about the natural world and the human role within it. For example, the environment is a mutable, moldable entity constantly changing in response to different magnitude forces, some subtle and others catastrophic. Change can occur in both positive and negative directions; thus, the effect of these forces may be to degrade or enhance nature or certain elements of natural systems. Last, the role of people and the ways in which we alter the structure of the environment and environmental processes can contribute to constructional changes in a positive direction or destructive changes in a negative direction.

Ecosystem restoration involves attempts to rehabilitate, fix, or heal once common plant communities, habitats that have disappeared or have become disconnected from one another, and the physical and chemical conditions necessary for their reestablishment and survival. It may include active biological manipulation employing techniques such as mowing, burning, planting, fertilizing, revegetation, the removal of nonnative or unwanted species, or the reintroduction of absent or nonviable species. Outside an explicit ecosystem focus, environmental restoration is aimed at the physical elements of landscapes such as watersheds and stream channels, hydrological function, or the removal and remediation of chemical pollutants. Other projects may be undertaken to reinforce unstable landforms to protect them from processes of erosion or mass movement; still others may be pursued with the goal of reestablishing a landscape aesthetic perceived to be more natural or, in the very least, more agreeable.

Historical Roots

Although the practice of environmental restoration may seem to be a very modern phenomenon, the roots of the concept run deep in North American geography. In the preface of Man and Nature, one of the most influential works in the history of American environmental thought, George Perkins Marsh expressed the hope that his book would begin to point the way toward understanding three fundamental points about the environmental function of the world around us. First, humans are a significant force at work changing the physical environment of the Earth. Second, many of these anthropogenic changes represent dangerous threats to the organic and inorganic worlds. Last, Marsh wanted to point out that people have the ability to act in the best interest of nature to redress past wrongs and enhance environmental conditions. He pointed to the importance of the restoration of disturbed harmonies and the material improvement of waste and exhausted regions. When Marsh heralded the significance of human agency in forcing environmental change and landscape degradation, he also provided reassurance that people could act as a force to
repair this damage. In a sense, Marsh is the originator of the modern concept of environmental restoration.

The rate of research investigations in environmental restoration has blossomed in recent decades. Prior to 1990, restoration was not a significant focus of explicit investigations, whereas currently, more than 1,000 articles are published each year that include keywords associated with environmental restoration across disciplines. This growth is seen by many to be symptomatic of the future of environmental management and applied environmental science. Like the biologist Edward Wilson, some have referred to the 21st century as the Era of Restoration or the Century of Restoring the Earth. Many of these investigations tend to be interdisciplinary, but specific subject area clusters where scholarly interest is highest can be identified. Areas in which the majority of academic research is being conducted include ecology, environmental science and toxicology, water resources, marine and freshwater biology, and forestry. Work in geography normally falls under one of these broad research umbrellas.

**Project Implementation and Contemporary Debates**

The scale of environmental restoration activities varies greatly from project to project. Large-scale plans such as the Florida Everglades ecosystem restoration project represent several complex mechanisms designed to redress a variety of ecological, physical, and hydrological changes that people intentionally made to the South Florida landscape in the past century. The enormity, complexity, and contentiousness of the plan has resulted in a project that will likely take decades to implement at a cost of tens of billions of dollars over an area spanning more than 46,000 km² (square kilometers). Others, such as the eradication of invasive species, for example, attempts to eradicate rats from the 26-km² Rat Island in the Alaskan Aleutian Island chain, small-dam removal efforts, or the localized remeandering of river channels, are more modest in scope and are aimed at changing the circumscribed elements of environmental systems.

One difficulty in describing the totality of environmental restoration is that it is often defined inconsistently and projects may have different goals attached to them. Some of this complexity derives from divergence over what aspect of the environment is to be restored. In 1988, the Society for Ecological Restoration was established in North America with the aim of re-creating indigenous, pre-European settlement conditions. Similarly, in 1992 the U.S. National Research Council defined the restoration of ecosystems as the return of an ecosystem to a close approximation of its condition prior to disturbance. Both descriptions reference a historical landscape, an ahistoric wilderness, of which little is known. In the United States, this historical reference to presettlement conditions and more recent variances such as ecosystem health or integrity is often synonymous with the concept of natural, an exceedingly complex term that implies a state where humans play no significant role in ecosystems, geomorphological and hydrological systems, or geochemical nutrient cycling. The human is antipodal to the natural. In this sense, what is being restored is the pristine landscape that existed prior to human manipulation and exploitation of environmental resources, a landscape separated from the imprint of people.

This has raised several concerns about both the practice and the concept of restoration among environmental ethicists. At issue is whether a restored environment is equivalent to the original one that was damaged or lost. In a seminal piece by Robert Elliot, restored environments are equated to art forgeries. They appear similar but do not have the same value of the original primarily because of the role of people in mimicking the natural processes forming these landscapes. Because humans and nature are defined as antithetical to one another, human activity is categorically different from natural activity and will inevitably produce artificial landscapes. As Eric Katz has noted, the result is an artifact. According to this line of reasoning, if the environmental degradation and despoiling that creates the need for restoration is due to human hubris and domination of the natural world, our attempts to actively repair the damage via the same means are yet another extension of this. The hubris is in thinking that it is possible
to repair the environmental damage and restore landscapes to their predisturbance value; the domination is in actively trying to make these changes come about by bypassing or manipulating natural processes for some intentional purpose. Whereas these debates have raised several ethical concerns about the concepts undergirding environmental restoration, they seem to have done little to slow the practical application of these concepts to environmental management.

The specific terminology used in environmental restoration strategies illustrates the subtleties in the intentions or goals of the projects. Restoration, as described above, is a transformation of environmental morphology and function to presettlement conditions. True restoration can be an impossible task because of the lack of detailed knowledge of prior states, the extent and magnitude of the biophysical changes that have occurred to landscapes, the impracticality of total removal of human elements from the system, and the lack of technical knowledge or resources needed to recreate relic environmental systems.

Restoration can also be viewed as a more generic umbrella term encompassing more specific terms such as reclamation, rehabilitation, and remediation, which maintain the presettlement environment as the target state but target more modest changes in the existing landscape. Reclamation can be viewed as a means of stabilizing highly degraded landscapes, re-creating conditions under which higher-order restoration activities such as rehabilitation can continue. Rehabilitation signifies a return of the environmental system to some discrete element of past structure or function that is either no longer present or severely degraded. Remediation indicates a focus on restoration activities related to the removal of pollutants or contaminants on sites significantly affected by human activities such as brownfields. All these terms employ “active” restoration strategies, whereby people actively take a role in reshaping the structure of the landscape or manipulating environmental processes to achieve the desired target state. Conversely, “passive” forms of restoration involve letting the environmental system heal itself through biotic regeneration, geomorphic adjustment, or chemical cycling. A critique of passive restoration is that it is too slow and assumes that the magnitude of degradation can be reversed without a permanent loss of essential elements of the system.

Other concepts broadly subsumed under the umbrella of environmental restoration break from the assumption that the target state should be presettlement nature or a historic environmental configuration. For example, naturalization has been used to describe the management of environmental systems to augment certain desirable functions through the manipulation of existing biophysical processes acting within the system. The current state of the landscape is inventoried and used as a baseline for subsequent change to determine progress and success. Whereas naturalization has been used mainly in geomorphic approaches to restoration, this emphasis on functional goals having the ability to self-regulate environmental structure by focusing on process mirrors recent discussions in restoration ecology as well.

Restoration is seen by many as an applied form of sciences such as ecology, toxicology, and geomorphology and as a design science, but one that must transcend scientific considerations. As such, concerns driving environmental restoration projects involve socioeconomic interests such as what people want from the environment, the elements of the system they value as important, and the tools or resources they are prepared to deploy, and these concerns determine the course of action as much as the constraints or techniques of systematic science. Thus, restoration is a cultural activity. In short, the process of determining what aspects of the environmental system are to be restored and how this will be accomplished is a political task, one that involves scientists but also environmental activists and the public.

Identifying success in environmental restoration projects is not always an easy or straightforward proposition. When specific restoration projects have ill-defined or generalized goals such as improvement, rehabilitation, or enhancement, the potential for these projects to be assessed for their effectiveness is limited. Another difficulty in assessing restoration accomplishments is that many projects are small-scale and fragmented from other small restoration projects and their effects are localized. Yet, as Jeffrey Bash and Clare Ryan noted, perhaps the greatest obstacle in the determination of restoration success is that
many of the projects are never assessed at all after completion. Whereas funds exist for the implementation of environmental restoration projects, it is much more difficult to obtain funding for monitoring activities. Although there is a trend of increasing spending in scientific research and public expenditure of funds on restoration activities of all kinds, this lack of monitoring and reliable determination of success represents a potential weakness in the ability of practitioners to learn from both their accomplishments and their failures. For environmental restoration strategies to continue being perceived as viable mechanisms of attenuating the deleterious human impacts on sensitive environments, the effectiveness and social desirability of scientific designs will need to be demonstrated as the field of environmental restoration matures.

Michael A. Urban

See also Brownfields; Environmental Ethics; Environmental History; Environmental Impact Assessment; Environmental Management; Environmental Planning; Environmental Services; Everglades Restoration; Forest Degradation; Game Ranching; Land Degradation; Landscape and Wildlife Conservation; Landscape Quality Assessment; Landscape Restoration; Marsh, George Perkins; Population and Land Degradation; Prairie Restoration; Soil Degradation

Further Readings


Environmental rights refer to the right to what are often referred to as “environmental services.” These include water, food, and materials for shelter. These services are intertwined with the evolving physical and cultural geographies of place. The geographic scope of such linkages has extended far beyond local influences to the planetary scale, and external pressures on both the quantity and quality of environmental services have created tensions and disparities that can be viewed by class, race, and gender at local, regional, and global scales. In this entry, the example of the right to clean water is highlighted as a case study. Environmental rights represent an area of intense debate in both academic and nonacademic circles and are a matter of life and death for the planet’s most vulnerable peoples. This is in large part because such rights lie at the intersection of traditional and “modern” resource management, technology, and economic development.
The environment has universally, in distinct ways, provided the basis for human existence, and formal and informal management regimes have evolved regarding access to it. Tensions emerge between groups regarding access to natural resources, and this stress is amplified when they are naturally, or made to be, “scarce.” At such times, the hierarchical structure of a given polity (or if global markets are involved, the interlinked capitalist system) may result in redistribution of resources according to the characteristics that determine power in that local system or overlapping multiple scales of systems. Such characteristics can include absolute or relative location, capital or class, connections to important political offices, gender, historical relations, linkages to other places or persons, military power, population size, or adjacency and political and economic linkages to powers outside the immediate region. Thus, environmental services become redistributed by the aforementioned systems at multiple scales to reflect these characteristics and power structures.

Injustices may occur in access to the quality or quantity of environmental services necessary for a certain group because they occupy an unfavorable position in the aforementioned hierarchies. The distribution of resources reflects the inequities in the system(s) they operate within. In addition, if the system extends beyond the local level, then it is not surprising when local people lose their environmental rights or have them stressed but are unable to work to improve their position. Thus, environmental rights are complicated by the multiscalar nature of a globalized world, the way we share it, and conflicting paradigms of development—and this can result in entrenched and systematic environmental injustice. Climate change is a high-profile example of this abstract nexus presently being played out from global to local scales.

Water

Arguably, the most well-documented recent battles regarding environmental rights have focused on the use of natural resources for development—the most well-known example being dams. The work of the World Commission on Dams illustrates this. However, in recent years, authors have increasingly discussed issues related to privatization, political and economic hierarchies, and the “human right to water” movement to address concerns about capitalism, justice, development models, and access to clean drinking water. Much of the well-publicized debate has been centered on the potential impact of free trade agreements on local governance of water, justice, and sustainability.

An important element of the discourse regarding justice and access to safe drinking water (and by necessity, sanitation) has been the rise of the human right to water movement and the need to fulfill “basic needs.” Global civil society has played a large role in advancing arguments that insist on essential levels of safe drinking water being provided regardless of ability to pay or any other characteristic. These advocates refer to such access as a “right” supported in UN documents. Many members of civil society have also insisted that there be a redirection in water development policies toward building local water management capacity through municipalities rather than the use of internationally privatized concessions as in Manila, where the largest privatization has largely failed. Table 1 provides a chronological look at policies developed to support the human right to water.

Environmental rights debates and conflicts are sometimes most starkly defined in the context of less wealthy places, but relevant examples also exist in industrial and relatively wealthy countries. For example, when water conservation is pursued, water conservation–oriented rates may be put in place. However, just as a “flat tax” has variable impacts on different economic levels of stakeholders, the implementation of water conservation structures also has differing impacts depending on the value of a unit of currency to an individual. One way to change the paradigm of water use to incentivize conservation for sustainability, but still protect low-income consumers’ finances and ability to access environmental services, is through the implementation of “inclined” block rate structures to provide “basic needs” levels of services at affordable levels (Figure 1).

As human populations and economic systems grow, and along with them consumption, the
**Type and Date** | **Contribution**
--- | ---
**Conferences**
1992 | The Dublin Statement on Water and Sustainable Development
1992 | Earth Summit in Rio, Agenda 21 (18.47), global action plan, chapter regarding freshwater
1996 | Heads of state and government meeting. Habitat II, Committee on Economic, Social, Cultural Rights
1997 | Comprehensive assessment of the world’s freshwater resources and discussion of “basic needs,” UN General Assembly Under Commission on Sustainable Development
1997/1998 | Sixth session of the UN Commission on Sustainable Development (CSD 6)
1998 | International Conference on Water and Sustainable Development Paris, France, UNESCO, and Int’l Institute on Sustainable Development (par 9)
2001 | Water Resources, Human Rights and Governance, Nepal
2001 | International Conference on Freshwater—aiming to build on the freshwater-related objectives identified in Chapter 18 of Agenda 21

**Constitutions**
1994 | South Africa Bill of Rights

**Covenants, major declarations, international agreements and resolutions**
1948 | Universal Declaration on Human Rights, UN
1966 | International Covenant on Economic, Social and Cultural Rights, UN
1966 | International Covenant on Civil and Political Rights, UN
1979 | Convention on the Elimination of All Forms of Discrimination Against Women, Art. 14.2(h)
1980, 1992 | “World Water Day” as passed in resolutions, UN
1980s | Declared to be the International Drinking Water Supply and Sanitation Decade, UN
1986 | Declaration on the Right to Development, UN
1989 | Convention of the Rights of the Child, UN (Art. 24)
1989 | European Convention on Human Rights, Council of Europe
1990 | New Delhi Statement, Global Consultation on Safe Water and Sanitation
1990 | World Summit for Children Plan of Action (par 5(d))
2000 | UN Millennium Declaration adopted by the UN Millennium Assembly
2001 | Committee of Ministers to Member States on the European Charter on Water (par 5 and 9 14)
2002 | “General Comment”; the UN Committee responsible for the Covenant on Economic, Social and Cultural Rights declares water a “human right”

**Sample rulings**
1985 | Inter-American court of human rights versus pollution of a tributary supply for the Yanomamis (Case 7615, Brazil)
2004 | An Indian state court (Kerala High Court) ruled against industrial overexploitation of a village’s groundwater
2008 | A South African court ruled against a prepaid water scheme in the town of Soweto on the basis of the basic right to water

**Table 1** Major legal contributions supporting the human right to water

*Source: Authors.*
demand for environmental services also grows. Apart from new technologies and new discoveries, the Earth has a stable or declining ability to provide the foundation for these “services,” and so debates and conflicts regarding environmental rights seem likely to intensify.

William James Smith Jr. and Ahmad Safi

See also Environmental Entitlements; Environmental Justice; Environmental Law; Environmental Racism; Environmental Security; Environmental Services; Human Rights, Geography and; Indigenous Environmental Practices; Water Needs

Further Readings


**ENVIRONMENTAL SECURITY**

The term *environmental security* suggests that environmental problems should be addressed as threats to human or national security. This can and has taken a number of very different forms, and as such there is no single, indubitable meaning of *environmental security*. It could include, for example, everything from environmentalists employing security rhetoric to draw attention to their political agendas to post–Cold War security officials appropriating new threats to justify exorbitant defense budgets. What an environmentally secure community, state, or world might look like is also debatable. This entry summarizes the ways in which scholars and institutions have conceptualized environmental security and examines debates and dividing points within the literature.

The current literature on environmental security mushroomed in the early 1990s, although the literature on environmental security has precedents predating the collapse of the Soviet Union. The Worldwatch Institute, along with several other environmentalists, argued repeatedly in the late 1970s and 1980s that national security was an outdated concept that needed to be revised in the face of global environmental change.

Environmental security’s key institutional moment, however, was the publication of *Our Common Future* by the World Commission on Environment and Development in 1987, which argued that pressing environmental issues and resource scarcities are best addressed through the security orientation that had thus far contained Soviet communism. This was a response to important policy debates about the environmentally disruptive effects of the rush toward “development” in the periphery and helped further instantiate environmental security into the foreign and development policy lexicon. The National Security Strategy statements produced annually by the Office of the U.S. president began incorporating environmental security regularly in 1991 (although the environmental content of these statements waned significantly 10 years later). Although environmental security was a prominent theme in the official policy statements of the Clinton administration, it is questionable to what degree and how it has ever been incorporated as a policy focus. For some of its proponents, environmental security entails state access and control of geopolitically significant resources—securing U.S. access to Caspian Sea oil, for example—although critics sometimes see environmental security as a pretense for neoliberal, capital-driven development policy.

In addition to official policy debates, there are a number of research programs designed to investigate the relationships (causative or correlative) between environmental change, scarcity, and violent conflict, perhaps most notably the Environmental Change and Acute Conflict Project at the University of Toronto. Such research programs are grounded in a case study approach to tracing out relationships between environmental...
degradation and violent conflict. Since its inception and increase in visibility, environmental security research in this form has faced (and responded to) a number of criticisms, the most prominent of which is that it typically analyzes only case studies in which the dependent variable, violence, occurs. Consequently, critics argue, they offer no control case, such as the existence of peaceful negotiations over a declining resource. Others stress the need to avoid explanations based on models of linear causality, in which independent variables (environmental change) are taken for granted as causes, rather than effects, of dependent variables. The more recent literature has sought to address these issues by focusing on, for instance, the role of natural resources (particularly transboundary water resources) in peace negotiations and pathways to environmental security focused on scales other than the state.

Beyond academic research programs, popular literature accounts of environmental security have been important as well. Another key moment for the environmental security literature was the 1994 publication of an article titled “The Coming Anarchy” in *Atlantic Monthly* by the journalist Robert Kaplan. Kaplan prophesied the spread of environmentally induced chaos from West Africa to other developing countries, and eventually to the developed world, and cited rapid population growth as the central cause. Kaplan’s thesis, and others like it, has at times been unfairly conflated with the research projects described above. It has also been the target of more recent environmental security scholarship, which is more critical of how the link between environment and conflict or security has been deployed and to what ends. Many such critiques come from the subfields of political ecology or critical geopolitics and argue that Kaplan’s neo-Malthusian bent is at some level characteristic of much environmental security scholarship. They also often eschew the construction of generalizable models of environmental conflict in favor of analyses focusing on the site-specific attributes of individual conflicts but situated within broader structures of power, such as international development policy, uneven terms of trade and lending, and consumption patterns in the developed world. Other critics also suggest the need to examine cases of resource abundance (such as diamonds or gold) and violence.

There exists little doubt that the environment and natural resources are important factors in the genesis of certain violent conflicts—the debate is about the nature of this relationship. Does environmental change drive such conflict, with local histories, culture, and political-economic structure at various scales merely the context that influences the form or nature of this conflict? Is the opposite the case, with environmental change something that exacerbates conflict—or for that matter is caused by conflict? Are these debates false constructs in the first place? Regardless, “environmental security” as a label does not have an unquestionable reference point in the real world; it means different things for different people in different contexts. For its various proponents, it can refer to state access to resources, the protection of the environment itself, the protection of people from the consequences of environmental change, or equal access to resources. In the latter case in particular, environmental justice has more recently become a key focal point for environmental security scholarship.

*Kolson Schlosser*

**See also** Critical Geopolitics; Ecological Justice; Environmental Entitlements; Environmental Impacts of War; Environmental Justice; Environment and Development; Malthusianism; Military Geography; Neo-Malthusianism; Political Ecology

---

**Further Readings**


Environmental services are processes and functions provided by the Earth’s myriad environments and ecosystems that sustain the human condition. The characterization, definition, mapping, and economic valuation of environmental services (also commonly referred to as ecosystem services) constitute a very active research area at the dawn of the 21st century.

One example of these services includes the pollination of crops and nonagricultural plants by bees and other pollinating species. The disappearance or extinction of bees alone would represent the loss of a nonmarketed ecosystem service that would have profound and painful impacts on the human condition. Another example is the erosion control services provided to various extents by vegetation. This example demonstrates how there can be a spatial dependence of the value of various ecosystem services. For example, a fire can dramatically diminish the erosion control services of a forest. The economic value of these erosion control services can vary based on proximity to capital infrastructure (e.g., reservoirs), agriculture, and vulnerable populations. A highly generalized list of ecosystem services would have to include the various regulatory, production, and control functions that ecosystems provide (Table 1).

The loss and/or degradation of various environmental services has the potential of inflicting significant social, financial, and environmental damage. Some of these damages will be market damages (negative impacts on gross domestic product [GDP]), and some are nonmarket damages (not accounted for in GDP). A clear example of a market damage that could be potentially attributed to a phenomenon such as global climate change is the economic value of the services provided by bees to pollinate crops and nonagricultural plants, which would have profound and painful impacts on the human condition.

### Table 1  Ecosystem services and ecosystem functions

<table>
<thead>
<tr>
<th>Ecosystem Service</th>
<th>Ecosystem Functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gas regulation</td>
<td>Regulation of atmospheric composition</td>
</tr>
<tr>
<td>Climate regulation</td>
<td>Regulation of global temperature, precipitation, and other climate processes</td>
</tr>
<tr>
<td>Disturbance regulation</td>
<td>Regulation related to ecosystem responses to environmental fluctuations</td>
</tr>
<tr>
<td>Water regulation</td>
<td>Regulation of hydrological flows</td>
</tr>
<tr>
<td>Water supply</td>
<td>Storage and retention of water</td>
</tr>
<tr>
<td>Erosion control and sediment retention</td>
<td>Retention of soil within an ecosystem</td>
</tr>
<tr>
<td>Soil formation</td>
<td>Soil formation processes</td>
</tr>
<tr>
<td>Nutrient cycling</td>
<td>Storage, internal cycling, processing, and acquisition of nutrients</td>
</tr>
<tr>
<td>Waste treatment</td>
<td>Recovery of nutrients and removal and breakdown of wastes</td>
</tr>
<tr>
<td>Pollination</td>
<td>Movement of floral gametes</td>
</tr>
<tr>
<td>Biological control</td>
<td>Trophic-dynamic regulations of populations</td>
</tr>
<tr>
<td>Refugia</td>
<td>Habitat for resident and transient populations</td>
</tr>
<tr>
<td>Food production</td>
<td>Fraction of gross primary productivity extractable as food</td>
</tr>
<tr>
<td>Raw materials</td>
<td>Fraction of gross primary productivity extractable as raw materials</td>
</tr>
<tr>
<td>Genetic resources</td>
<td>Sources of unique biological materials and products</td>
</tr>
<tr>
<td>Recreation</td>
<td>Providing opportunities of recreational activities</td>
</tr>
<tr>
<td>Cultural</td>
<td>Providing opportunities of noncommercial uses</td>
</tr>
</tbody>
</table>

change is the loss of tourism dollars associated with visits to coral reefs that are destroyed by sea-level rise and or sea temperature change. An example of nonmarket damage is the loss of the ecosystem service of storm protection provided by coastal wetlands in a sea-level-rise scenario. Perhaps ironically, the value of the storm protection services of coastal wetlands (on a per-hectare basis) actually goes up with climate change–induced increases in storm frequency and intensity. Traditional economic paradigms are increasingly challenged by environmental problems, and assessing the costs and benefits of changes to ecosystem services from a strictly economic perspective is particularly problematic.

1. Human development has actually increased many ecosystem services; however, it has also significantly damaged or destroyed many others.

2. UN plans to eradicate famine and disease are significantly impeded because of extant and probable damage to ecosystem services around the world.

3. Human demands for food, fresh water, timber, fiber, and fuel have been the driving force of irreversible losses of biodiversity and ecosystem degradation.

4. Facing these challenges effectively will require the establishment and/or creation of policies, institutions, and practices that do not currently exist.

Economic valuation of ecosystem services (whether they are regarded as public goods or common property) is one rhetorical means of clarifying the importance of both ecosystem services as they stand and the consequences of their continued degradation. It is reasonable to argue that the total value of the world’s ecosystem services is infinite or incalculable simply because humanity could not exist without these services; however, most attempts to place a dollar value on ecosystem services are made through economic valuations at the margin (e.g., What would the loss of the ecosystem services provided by 1 square kilometer of wetlands cost?). Attempts to make these kinds of global assessments of marginal values nonetheless produce global values of ecosystem services in the magnitude of tens of trillions of dollars per year. These likely underestimates of the total value of the world’s environmental services nonetheless exceed the total value of the world’s marketed economy. We ignore the value of environmental services at our own peril.

Paul C. Sutton

See also Common Pool Resources; Ecological Economics; Ecotourism; Environmental Entitlements; Environmental Security; Global Sea-Level Rise; Market-Based Environmental Regulation; Neoliberal Environmental Policy; Oceans; United Nations; United Nations Environment Programme (UNEP); Wetlands
The concept of environment and development is an attempt to integrate economic development policy-making and environmental considerations. It encompasses an acknowledgment that millions of people worldwide still live in poverty; that there is continued unequal access to the resources needed for a decent livelihood; and that there is a need for economic growth and that many human activities, including the quest for survival, put a strain on the environment. Since the early 1970s, environment and development has gradually gained momentum as an organizing concept for debates and policies on economic development and environmental conservation. It has found its way into the language of international bodies and agreements addressing environment and development, the policies and legislation of many countries, subfield or related individual courses of study in academic institutions, and academic journal titles or focus areas, as well as becoming a convenient way to link disparate groups in the development-environment nexus.

Like many other concepts within the development discourse (e.g., development, sustainability, participation), the environment and development concept is both loaded and slippery, allowing for multiple interpretations among its users. Perhaps owing in part to the ambivalent implications of the root words that make it up, the concept of environment and development conveys both optimism and a warning. It is optimistic in the sense that it conveys an idea that there can be positive links between economic growth and environmental conservation as long as enabling policies are in place. The concept also carries a warning that the natural resource base is not infinite and that if misguided policies are pursued, economic growth through industrialization and overexploitation or misuse of natural resources could lead to environmental disasters that can negatively interfere with life as we know it today.

Like the concept of sustainable development, environment and development also has been viewed as an oxymoron in that there is an implicit goal of economic growth, but it is to be achieved while remaining within given ecological parameters. Wolfgang Sachs has even characterized the link between environment and development as a “dangerous liaison,” in the sense that, again like sustainable development, the balance is never easy to achieve or even measure.
global thinking about this concept through their sponsorship of activities where it has received significant attention. Some of these key institutional foundations of environment and development are discussed below.

### UN Conference on the Human Environment

The UN Conference on the Human Environment, held in Stockholm in June 1972, can be considered as having laid the foundation for the linkages between environment and development debates and policies on an international stage. First proposed by Sweden, which was concerned about acid rain and industrial pollution, the Stockholm conference’s main focus was to forge international cooperation for and on the environment. It is in fact considered as the occasion on which the international community’s attention on environmental concerns received serious attention and provided momentum for their ascent to the top of the international agenda. Most significantly, the conference emphasized that environmental problems were caused not only by industrialization but also by poverty and lack of economic development. Thus, development, particularly poverty reduction, was needed for safeguarding the environment. Other significant outcomes of this conference include the establishment of the influential United Nations Environment Programme (UNEP) in Nairobi, the establishment of environmental ministries and agencies in more than 100 countries, and a significant increase in the number of nongovernment organizations dedicated to environmental conservation.

The link between environment and development was further clarified in a follow-up to the Stockholm conference, the Symposium on Patterns of Resource Use, Environment and Development Strategies, held in Cocoyoc, Mexico, in 1974. This symposium emphasized the need to see the development and environment problem as also being rooted in the unequal distribution and misuse of natural resources, rather than scarcity alone. However, despite the gains made by the Stockholm conference and its follow-up, environmental problems continued. Nevertheless, the successes of this conference and the continuing environmental problems appear to have strengthened the resolve of international agencies to work toward a solution.

### The World Conservation Strategy

In 1980, UNEP, one of the notable products of the 1972 Stockholm conference, in collaboration with the International Union for Conservation of Nature and the World Wildlife Fund, produced the World Conservation Strategy. As part of this strategy, the concept of sustainable development was introduced on the international stage. Sustainable development was presented as an attempt to bridge the gap between environmental concerns about ecologically destructive human actions and sociopolitical concerns about human development issues. Significantly, the World Conservation Strategy identified the maintenance of essential ecological processes and life support systems, the preservation of genetic diversity, and the sustainable use of species and ecosystems as the three primary goals of conservation. Even though this strategy had roots in wildlife conservation, it was the first comprehensive strategy to link conservation and sustainable development.

### The World Commission on Environment and Development

Whereas the World Conservation Strategy introduced the concept of sustainable development into the international agenda, the World Commission on Environment and Development (WCED) is responsible for popularizing it. This commission produced its report, titled *Our Common Future*, or the Brundtland Report (after the commission’s chair, Gro Harlem Brundtland, the former prime minister of Norway), in 1987. *Our Common Future* is arguably the publication that provided institutional legitimacy to international concerns and hopes concerning environment and development. Christopher Barrow has even argued that the Brundtland Report can be singled out as a landmark text of the 1980s and 1990s. Although it has been challenged and modified the world over, the definition of sustainable development that *Our Common Future* offered—economic development that meets the needs of
the present without compromising the ability of future generations to meet their own needs—remains a benchmark in almost all discussions of this issue.

The WCED’s main goals were to reexamine critical environment and development problems and to formulate realistic proposals for solutions. Thus, the critical objectives for environment and development policies that were put forward in Our Common Future include reviving economic growth; changing the quality of growth; meeting essential needs for jobs, food, energy, water, and sanitation; ensuring sustainable levels of population; conserving and enhancing the resource base; reorienting technology and managing risk; and merging environment and economics in decision making. Poverty eradication was yet again seen as a desirable outcome of a successful development process and a prerequisite for environmental protection.

The World Bank’s World Development Report

Perhaps a less publicized, yet significant milestone in the institutional history of the environment and development concept is the World Bank’s World Development Report 1992: Development and the Environment. Although its main message of the need to integrate environmental considerations with development policy making does not differ from those of the Stockholm Conference, the World Conservation Strategy, and the Brundtland Report, the World Bank’s apparent endorsement of those earlier observations and goals concerning environment and development is quite significant.

One significant variation between the World Bank’s report and the earlier ones was the Bank’s primary focus on poverty and developing countries. It emphasized the difference between the environmental problems facing the poor and those of wealthier countries, arguing that whereas industrialized countries need to solve their own problems, they have a role to play in helping improve the environments of developing countries. In its suggestions, the report highlighted the clear links between the environmental problems of developing countries and the economy and lifestyle activities of wealthy nations. Besides the 1992 report, the World Bank’s subscription to the environment and development concept saw it establish an environment department and regional environment divisions as early as 1987. In the World Development Report 1992, the Bank articulated its commitment to have environmental considerations more deeply embedded in every aspect of its work.

The UN Conference on Environment and Development

One of the main suggestions contained in the Brundtland Report of 1987 was that the UN should organize an international conference to review progress and promote new initiatives concerning the environment and development agenda. Thus, in June 1992, more than 170 governments, numerous heads of state, more than 1,000 nongovernmental organizations (NGOs), and thousands of other participants and campaign groups assembled in Rio de Janeiro, Brazil, for the UN Conference on Environment and Development, sometimes called the Earth Summit. Of the five key outputs of this conference, which included conventions (treaties) and declarations on forests, climate change, and biodiversity conservation, two specifically relating to environment and development can be singled out. The first one is the Rio Declaration on Environment and Development (the Earth Charter), which is a statement of 27 key principles to guide the integration of environment and development policies. These included statements on the right of every country to have its own policy on the use of its natural resources; the right to development; the necessity to reduce consumption patterns that conflict with the goal of sustainable development; and many others. The second and perhaps most significant output of the conference is Agenda 21 (a plan for the 21st century), which is a plan of action to help with sustainable development efforts following the Earth Summit. In its 40 chapters, the Agenda 21 document sets out objectives and priorities, identifies institutional questions, and assesses the costs of the proposed measures. Despite its ratification by most of the nations present, there were doubts
about whether the plans of Agenda 21 were financially feasible. Nevertheless, Agenda 21 raised much hope, especially because it was envisaged as a more community-led process that relied on local initiatives by governments and people.

In addition to stressing the importance of partnerships between countries, as well as between governments and citizen groups within each nation, the WSSD had two significant products relating to environment and development. The first one is the Johannesburg Declaration on Sustainable Development. This political document, which contains 37 articles outlining the path taken from Rio to the WSSD, highlights present challenges, expresses commitment to sustainable development, underscores the importance of multilateralism, and emphasizes the need for implementation. The second significant output of the WSSD is the Johannesburg Plan of Implementation for Sustainable Development. Contained in 10 chapters, the plan is a framework for action to implement the commitments that were agreed on in Rio. There are, arguably, three notable chapters relevant to this discussion. Chapter II is on poverty eradication and is based on goal number one of the eight Millennium Development Goals—of halving the proportion of the world’s poor by the year 2015. This chapter also attempted to clarify the complex links between poverty and sustainable development. Chapter IV is on the protection and management of the natural resource base of economic and social development, including action on water resources, natural disaster mitigation, and sustainable use of minerals, metals, and mines, to mention a few. Chapter VIII focuses on sustainable development for Africa. It highlights poverty, violent conflicts, inadequate investment, and human immunodeficiency virus/acquired immune deficiency syndrome as the factors hindering sustainable development in Africa. It offers several recommendations for pursuing sustainable development in Africa, including sustained economic growth, reversing desertification, integrated water management, sustainable agricultural production, strengthened health care systems, and sustainable tourism, among others.

A review of the conferences and reports on environment and development of the past four
decades easily conveys the notion that this concept is easier to talk about than to put into practice. The years between the Stockholm conference, in 1972, and the WSSD, in 2002, are characterized by what observers see as pedestrian progress on the goal of linking environmental considerations and economic development. There are likely many reasons for this slow pace of change, but only four possibilities are discussed here.

The first challenge for integrating environment and development is that both environmental conservation and development are, individually, characterized by conflicts over meaning, particularly the “how” aspect, as well as about who is responsible. Bringing the two concerns together therefore presents immediate challenges; nevertheless, it is possible to overcome these challenges. It is perhaps safe to acknowledge that linkages between environment and development are dynamic and context specific, reflecting geography, scale, and social and political issues among the groups involved.

The second challenge for integrating environment and development is that even though some countries have single ministries that address environmental issues and development fall under different departments. Facing political and financial pressures, individual ministries are often forced to turn a blind eye to commitments to integrate the two issues.

The third challenge is that given the massive global inequalities and differences in political landscapes between nations, international cooperation on environment and development is not unique in the challenges it faces. Both wealthy and poor countries individually have their own priorities and challenges.

Last, a review of the institutional genealogy of environment and development leaves a strong impression that its goals and those of sustainable development are fairly similar. In that case, one of the major critiques of sustainable development, which is that it is an oxymoron—seeking economic growth but staying within ecological limits—likely applies to the environment and development concept. However, like sustainable development, the environment and development concept has survived for more than three decades and continues to motivate dialogue, which in turn always has possibilities of leading to positive action.

Thembela Kepe

See also Environmental Entitlements; Environmental Ethics; Environmental Impacts of Cities; Environmental Impacts of Manufacturing; Environmental Protection; Environmental Security; Environmental Services; Malthusianism; Nature-Society Theory; Neoliberal Environmental Policy; Neo-Malthusianism; Polychlorinated Biphenyls (PCBs); Population and Land Degradation; Population and Land Use; Population, Environment, and Development; Sustainable Development; Sustainable Development Alternatives; Sustainable Fisheries; Sustainable Forestry; Sustainable Production; United Nations Conference on Environment and Development; United Nations Environment Programme (UNEP); World Bank; World Summit on Sustainable Development

Further Readings

Knowledge is a type of belief, and epistemology is the study of knowledge, that is, the analysis of how we know the world, as opposed to ontology, which deals with what we take to be real. An epistemology is a worldview with a particular set of assumptions, a definition of truth, and priorities concerning what constitutes valid and important knowledge. Knowledge comprises those beliefs that are both true (they conform to what is real) and justified (the one who thinks them true has sufficient reason to think them true). This formulation attempts to formally distinguish the type of belief we call knowledge from other types of belief, such as errors (untrue beliefs) and lucky guesses (true beliefs held without justification).

Everyone has an epistemology, whether they know it or not. In academic analysis, epistemologies are made explicit as they underlie the rules of knowledge formation. Different epistemologies have different criteria for what constitutes valid knowledge, different definitions of truth, and different priorities in terms of how analysis should be done. Each piece of knowledge can be considered a proposition. Propositions can be sorted (not without remainder) into three epistemic classes: demonstrable propositions (those that can be shown to be true by appeal to logic or fact), defensible propositions (those for which a convincing rational argument can be made), and preferred propositions (those that reflect individual subjective taste and cannot or need not be defended). We hope to win another person’s assent to a demonstrable proposition when we can “show” them it is true in a manner that all rational humans will find persuasive. The truth of such statements is given either a posteriori, or as empirically evident to the senses (e.g., “Water runs downhill”), or a priori, that is, by definition and logical deduction (e.g., “A man cannot marry his own widow”).

Other propositions may not be clearly demonstrable on empirical or logical grounds, but arguments can be offered in their favor. Dispute over defensible propositions results in agreement only some of the time. This is because their truth or falsehood cannot be conclusively demonstrated beyond the shadow of reasonable doubt. Disputants can only advance arguments for and against such propositions. Assent (or dissent) follows a judgment, or weighing, of arguments, reasons, and considerations, but unanimity may not be attained. An example of such a proposition is “The death penalty is unjust.” Each person believes that the judgment with which she or he responds to a question such as “Is the death penalty justified?” is true and justified but, in the course of dispute with those with a different answer, finds that his or her judgment is, at best, defensible and cannot be “proven” empirically or logically. No one person’s argument can carry the day, for other people’s judgments are defensible too.

Last, some propositions state preferences that are simply subjective expressions of individual taste. “Wensleydale is the finest cheese” or “Fall is the finest season” are propositions of this sort. In asserting them, people have no empirical evidence or inescapable logic to advance in their favor. Because they simply express a preference, they require no defense. No one knows what I like better than I, thus the famous adage: de gustibus non est disputandum (“There is no disputing taste”).

Epistemology and Geography

Several epistemologies have shaped geography over time, each rooted in a different conception of truth, valid knowledge, and rules of argument. The word epistemology existed in English for more than a century before geographers used it. In the 1960s, the term provoked controversies when it was deployed by a new generation of epistemological puritans, known as logical positivists, who demanded very strict standards of justification (or verification).

Logical positivism is the insistence that only demonstrated propositions (those based on logic or facts) should count as knowledge. It demotes all defensible propositions (those based on rational argument) to mere subjective preferences. For positivists, facts are the most important and incontrovertible form of evidence used to sustain a proposition. The real significance of positivism, though, is that it seeks to separate propositions into just two epistemic classes: (1) demonstrable propositions, which, once demonstrated, cannot be reasonably disputed, and (2) mere expressions of preference. For positivists, only knowledge
grounded in logics and facts counts as “true.” One rub, however, is that positivist claims are not themselves demonstrable propositions but rather are judgments open to dispute.

Post-positivist epistemology developed outside geography, including philosophy and literary criticism. There are a variety of post-positivist schools of thought, including Marxism, humanistic approaches, and various postmodernist or poststructuralist interpretations, which share certain broad features in common. This set of perspectives typically recognizes two things: (1) that every knowledge claim rests on presuppositions that are, at best, judgments and (2) that within the realm of judgment, there is more than one style of valid reasoning (i.e., verification procedure) that humans use and find persuasive. An empiricist cannot, for instance, empirically verify empiricism without begging the question, and theological, historical, and critical arguments employ verification procedures different from (and less conclusive than) those used in science. These considerations lead many to surmise that there are multiple rationalities, each with its own “logic,” or what many today call plural epistemologies. Post-positivist epistemologies tend to focus on knowledge as a social construction, linked to particular interests, with social consequences. Much of this theorizing concerns the origins, nature, and impacts of discursive representations of the world.

Engaged in dispute, each of us attempts to promote our own beliefs while demoting the beliefs of our opponents. We typically present our defenses as demonstrations while treating our opponents’ judgments as preferences or as a mere question of taste. Epistemology in geography is, in short, very often a rhetorical tool employed in the hope that our beliefs will be taken more seriously than those we dispute.

Jonathan M. Smith

See also Discourse and Geography; Ethics, Geography and; Existentialism and Geography; Geographical Imagination; Human Geography, History of; Humanistic Geography; Idiographic; Kant, Immanuel; Logical Positivism; Nomothetic; Ontology; Phenomenology; Positionality; Postmodernism; Poststructuralism; Realism; Situated Knowledge; Structuralism

Further Readings


The equator is the great circle farthest from the north-south spin axis of Earth. The equator is everywhere equidistant from the North and South poles on the surface of Earth, which is a sphere or a reference ellipsoid.

Although the equator is easy to define, it is not easily found and marked on Earth. The gravitational force is lower along the equator since it is the circle on the ellipsoid farthest from the center of mass of Earth. The equator marks the line above which the Coriolis force deflects movement toward the left of the direction of travel and below which the deflection is to the right. While these effects are real and influence weather patterns, ocean currents, and aircraft flight, they are so weak as to be almost immeasurable at the equator. Polaris, the pole-star, can be used to find zero latitude, but atmospheric effects make it difficult to observe accurately at the equator.

However the equator is defined by any of more than 100 geodetic datums in use in the world, it is a physically definable origin for latitude. The geodetic equator is the circle on the surface of a reference ellipsoid equidistant from the poles of rotation of the ellipsoid. For different ellipsoids, the geodetic equator can be in different places on the physical surface of Earth.

Since there are various definitions for other latitudes, there are other equators found along the line representing 0° of those latitudes. There is a celestial equator, defined with respect to celestial latitude, an astronomic equator, a galactic equator, a geomagnetic equator, and others.

The equator has meanings in many contexts. The geostationary weather and communications
satellites orbit Earth above the equator. The tropical sun’s rays are at their highest at midday along the line. Along its course, tropical rain forests, massive river systems, and sparse populations are found on land. At sea, sailing ships can be trapped in the equatorial doldrums, and sailors still perform ceremonies for those crossing the line for the first time. Joseph Conrad, Mark Twain, and Herman Melville have used the equator in the titles of their books. There have been over a dozen films with “equator” in the title. There is a region known as Equateur in the Republic of Congo, the Equatorial Channel in the Indian Ocean, the countries of Equatorial Guinea and Ecuador, and the group of Pacific islands known as the Line Islands or the Equatorial Islands.

There are dozens of equatorial monuments in South America, Africa, and Asia, but they are often tourist facilities more than they are accurate markers of the location of the zero for the variously defined systems of latitude. For example, the Mitad del Mundo monument in Ecuador marks the approximate position of zero latitude for a system no longer in use in Ecuador or anywhere else in the world. On Ecuadorian topographic maps, the monument is responsibly placed several hundred meters above the line representing zero latitude for the Provisional South American Datum of 1956. GPS-equipped visitors to the Middle of the World will find the World Geodetic System of 1984 zero latitude a few hundred meters north of the monument.

Peter H. Dana

See also Datums; Earth’s Coordinate Grid; Geodesy; Latitude; Longitude; Poles, North and South

Further Readings


The term *equinox* (in Latin *aequinoctium*) derives from two Latin words: *aequus*, whose meaning is “equal,” and *noctium* (*nox, noctis*), referring to “night.” Its definition involves that of the celestial sphere (which is the infinite apparent sphere around the Earth that is imagined to have our planet as its center and that contains the sun, the moon, and all the planets and stars), the ecliptic (the plane on which the orbit of the Earth around the sun lies), and the celestial equator (the huge imaginary circle on the celestial sphere that lies in the same plane as our planet’s equator): An equinox, in fact, can be explained as either of two points (equinoctial points) on the celestial sphere where an intersection between the ecliptic and the celestial equator occurs. On the other hand, it can also be defined as either of the two yearly times when the sun crosses the celestial equator. At those specific moments, the sun rises exactly east and sets exactly west, and as can be inferred from the etymology of the word underlined at the beginning, the length of the day (sunlight) and that of the night (darkness) are nearly the same, irrespective of latitude. Vernal equinox is around March 20, and it marks the beginning of spring in the Northern Hemisphere, whereas autumnal equinox is around September 22, and it is the beginning of the cold season. Just the opposite occurs in the Southern Hemisphere. From an anthropological point of view, extremely interesting are those myths and rituals that refer to the equinoxes and are related to ancient cultures all over the world.

Equinoxes are said to have been discovered in ancient times by the Milesian Anaximander, who applied the gnomon (a style fixed on a horizontal plane that projects a shadow) that allowed him to determine them together with the solstices and the inclination of the ecliptic; a few centuries later, the Grecian astronomer and geographer Hipparchus of Nicaea was credited with the discovery of the precession of the equinoxes, which he studied by comparing the results of his observations of the sky with those of his predecessors. He stated that equinoxes are not always the same as they move westward along the ecliptic if compared with the assumed fixed stars on the celestial sphere. This phenomenon is due to the slow movement
(precession) of Earth’s rotational axis caused by the gravitational forces of the sun and the moon on the equatorial bulge of our planet. Considering that precession is approximately $1^\circ$ in 72 yrs. (years), the equinox line moves clockwise and completes its $360^\circ$ round in 25,800 yrs. As a consequence, every year the dates of equinoxes come slightly earlier.

_Susanna Servello_

See also Anaximander; Equator; Hipparchus; Latitude; Solstices

Further Readings


**ERATOSTHENES (276–194 BC)**

Eratosthenes of Cyrene (born in Cyrene 276 BC, died in Alexandria 194 BC) was the first scholar to use the term geography, meaning by that the description of Earth. His studies pertained to geography, astronomy, and mathematics and allowed him to confirm the hypothesis about the round shape of the planet and to measure its size with an extremely good approximation. He was also the first to imagine dividing Earth’s surface into portions demarcated by meridians and parallels. With a deep interest in nature, he studied tides and their lunar cycle, and by observing the presence of marine fossils far from the sea, he realized the slow transformation of the coastlines.

The starting point that made Eratosthenes estimate the measurement of Earth’s circumference was the observation that at noon, on the summer solstice, the bottom of wells located in the Egyptian city of Syene were vertically illuminated by the sun, while at the same time in the northern city of Alexandria, objects (in particular he was observing an obelisk) were casting their shadows. After noticing that in Syene the sun was at the zenith (he considered it to be crossed by the Tropic of Cancer) and presupposing that the two cities were located on the same meridian, he calculated that the distance between them was around 5,000 stadia, based on the strength of the observed average speed of the caravans of dromedaries. He also correctly asserted that due to the enormous distance that separates the sun from Earth, its rays could be considered parallel when they reach the surface of the planet. Considering that in Alexandria the beams were not perpendicular to the surface but were inclined at an angle of $7.2^\circ$ relative to the vertical direction, through trigonometry he got the result that the angular distance between the two cities was $1/50$ of Earth’s circumference, which let him calculate both its dimension and its distance from the sun. He also succeeded in estimating the inclination of the ecliptic.

As a cartographer, Eratosthenes mapped the river Nile from Khartoum to its delta. He also drew a map of the known world, with Persia at its center and its limits described as follows: Thule (the island, described for the first time by Pytheas, that could be equated with Iceland) in the north, Ethiopia in the south, the Pillars of Hercules in the east, and India in the west. The Caspian Sea was imagined to communicate with the Northern Ocean, and as far as the African continent is concerned, its whole northern part was named Libya. Eratosthenes is also supposed to be the inventor of the armillary sphere (also known as spherical astrolabe), which is a model of the celestial sphere.

_Susanna Servello_

See also Anaximander; Aristotle; Cartography, History of; Equator; Hipparchus; Human Geography, History of; Latitude; Longitude; Ptolemy; Solstices; Strabo

Further Readings

**ERROR PROPAGATION**

Quantitative attribute and geocoded data collection almost certainly entails error that arises from five sources: calculation, measurement, specification, sampling, and stochastic processes. Error propagation may be defined as the transmission of this error from data inputs to output results when performing data manipulations on maps that possess such an error. In extreme cases, this situation is characterized by the maxim “garbage in, garbage out.”

### Types of Error

- **Calculation error** refers to mistakes that arise from incorrect executions of arithmetic operations with, or imprecise representations (e.g., rounding) of, numbers.
- **Measurement error** refers to incorrect numbers in a data set. It arises from the use of the wrong measurement scale (i.e., nominal, ordinal, interval, ratio) or instruments lacking precision (i.e., the number of digits to the right of a decimal point), and/or recording mistakes.
- **Specification error** refers to the use of incorrect assumptions and/or the application of incorrect mathematical formulae or equations in a data analysis.
- **Sampling error** refers to differences that are due solely to the use of a subset (a sample) rather than an entire collection of objects (a parent population). Finally, **random noise** in a geographic landscape represents a haphazard (stochastic) disturbance and is characteristic of ignorance or uncertainty about reality or is the outcome of a goodly number of factors that jointly behave in such a way that random disturbances appear to be introduced into attribute and/or geocode measures.

### Controlling Error Propagation

Georeferenced data error rarely averages to zero during data manipulations. Rather, error propagation can be controlled by minimizing or eliminating data input and/or analysis error. The widespread use of computers and hand calculators has minimized calculation error, which today tends to be associated almost solely with rounding and the number of significant digits to the right of a decimal point. Improvements in instruments (e.g., decreasing the pixel size in satellite remote sensors, global positioning system [GPS] technology) and the digital rather than manual handling of data, for example, have reduced measurement error. In particular, the development of spatial statistical theory, statistical theory for non-normal data, and resampling techniques has helped diminish the occurrence of specification error, which continues to be the principal source of error propagation. Experimental design and statistical mixed-model approaches (i.e., including random effects terms) furnish controls for sampling error, which because of its random nature is very controllable. Random noise in a geographic landscape refers to model residuals (i.e., the difference between predicted and observed values) and can be controlled by acquiring a better substantive understanding of geographic phenomena.

### Relative Importance of the Origin of Error

Error propagation with georeferenced data is complicated by the presence of pairwise attribute correlation as well as positive spatial autocorrelation (i.e., attribute and error values at adjacent locations tend to be similar). Propagation occurs when these data are manipulated within a geographic information system (e.g., overlay, buffering) or used to construct indexes (e.g., linear combinations or ratios of, say, spectral bands). For univariate overlay, overlay-AND, overlay-OR, and overlay-XOR, attribute error and location error tend to dominate error propagation, with spatial autocorrelation of attribute variables often being the single most important contributor to autocorrelation in propagated error. Attribute error tends to be the single most important contributor to error propagation when addition or ratioing is performed, with interattribute
correlation playing a more important role for addition and location error playing a more important role for ratioing. Synchronized attribute and error map patterns frequently obscure the detection of propagated error.

In a multivariate context, either high interattribute correlation or high positive spatial autocorrelation often moderates the overall effects of error propagation. However, summarizing a set of highly correlated variables with only the first principal component, which captures much of the variability in a given data set, tends to exacerbate error propagation.

**Digitizing Error**

Because so much geospatial information has been digitized manually over the years, researchers have studied the positional error (i.e., the discrepancies between reference and recorded locations) associated with this digitizing process for several decades, finding, for example, that it can be characterized by a bivariate normal distribution.

Today geocodes are typically obtained with global positioning systems or address-matching software. Although these technologies have improved the locational accuracy of georeferenced data, they are not error free. This positional error affects the results of any spatial statistical analysis performed with a georeferenced data set. One outcome of error propagation is the recurrence of spatial clustering in positional error at various geographic resolutions (i.e., aggregations). One empirical study suggests another outcome, namely, the identification of a noticeable but not shockingly large impact from positional error propagation in spatial regression analysis results for the data set analyzed.

**A Critique**

Error enters into a geographic analysis in many ways and is usually propagated in an unadulterating or amplifying, rather than a diminishing, fashion. Besides the multitude of error propagation possibilities associated with attribute data alone, additional complications are introduced by the presence of spatial dependencies and by positional error. This error propagation can have serious consequences on decision making or knowledge creation. For the most part, computational error is being managed well. Considerable efforts and investment are being made to manage measurement error, with GPS technology helping reduce positional error. Perhaps the single most problematic feature of this error source concerns confidentiality and the frequent need to geographically aggregate data. Spatial statistics and spatial econometrics, as well as modern advances in statistical methodology (e.g., the generalized linear model, Bayesian map analysis, mixed models), are contributing to better management of specification error. Spatial sampling designs, such as tessellation-stratified random sampling, devised for the Environmental Monitoring and Assessment Program, are enabling more effective management of sampling error for a georeferenced data collection. Random noise in a geographic landscape is, perhaps, the most difficult error source to manage because its management is at the mercy of scientific discovery.

*Daniel A. Griffith*

See also Digitizing; Ecological Fallacy; Ground Reference Data; Interoperability and Spatial Data Standards; Metadata; Modifiable Areal Unit Problem; Ontological Foundations of Geographical Data; Spatial Data Integration; Spatial Data Mining; Spatial Data Models; Spatial Data Structures; Spatial Econometrics; Spatial Interpolation; Spatial Multicriteria Evaluation; Spatial Resolution

**Further Readings**


Ethics is an inquiry into the moral values embodied in discourse and practice and a concern for what is good, right, or just in our individual and collective lives. It is an attempt to formulate rules of thumb to help us grasp the ends and means of life, providing insight and guidelines to strive for what the ancient Greeks termed *eudaimonia*, a term sometimes translated as happiness but better understood as “flourishing.” The ancient geographer Strabo referenced this notion when he noted that geographers and ethicists alike are interested in “the art of life, that is, of happiness.”

Ethics can be a subject that is difficult to discuss in geography (and elsewhere) for it raises fears of dogmatic worldviews. There are indeed people who use ethics to scold others, score debating points, or justify doctrinaire approaches to life. But this is not the main tradition of ethics. Rather, as Socrates notes in Plato’s *Republic*, ethics is an exploration of “how we ought to live.” It is a conversation about the moral values that inform (or ought to inform) our way of life. This involves a process of critique and vision. We criticize what detracts from our well-being, and at the same time, we envision how we might improve our lives.

Ethics may be informed or distorted by religion, spirituality, personal experience, or social custom, but it is not reducible to these sources. Instead, it is a reasoned and evidentiary dialogue that bridges cultural and disciplinary positions to improve the well-being of ourselves and others. These others can include different entities (e.g., human or nonhuman) considered at different scales (e.g., local to global, individual to system). Thus, ethics may concern itself with the well-being of people, animals, and the rest of nature, whether they present themselves as individuals or communities, ecological systems or societies, over space or through time.

**Ethics and Social Change**

Ethics is also a form of discursive power. It helps reveal moral concerns, guide our thoughts and actions in addressing moral problems, and hold people and societies accountable for their actions. Moral critique and vision are the foundation for all movements of social change, whether these are for animal, environmental, or social causes. It is for this reason that ethics is indispensable in political life generally, as well as in the life of the academy. This is not to say that the moral norms embedded in social customs and laws are always or mostly right. We need only look at the transformation of norms regarding race, gender, and sexual identity for examples of moral progress. Even so, ethics-based arguments motivate struggles for change and spur the evolution of our customs and laws.

If this arc of ethics and social change seems a crooked path, think of it as akin to the development of law, medicine, or any practice-based tradition of knowledge. There is much wrangling, and there are many errors, but over time, trends emerge that point toward better ways of engaging the world. Reason and evidence are key here. They can do much to contest invidious custom and prejudice. They also help adjust our moral compass to distinguish better from worse norms and practices.

**Ethics and Geography**

It may come as a surprise to some that geography has a strong streak of ethics in its discourse. During the quantitative revolution and the hegemony of logical positivism, ethics was peripheralized by the theoretical and methodological dogmas of a putatively value-free and ethically neutral scientism. With the steady erosion of scientism in geography, however, ethics has been revitalized as a living tradition of geographic thought, and geographers are experiencing a moral turn in their research and practice.

This moral turn began several decades ago with humanists, Marxists, feminists, environmentalists, and others seeking greater engagement with the social and environmental issues of their time. Many geographers began to investigate ethics-laden questions of research practices, protection of human subjects, cultural diversity, social justice, environmental protection, and the like as part of their research. Moreover, they began to speak with a moral voice about what ought to be done to allow people and the planet to flourish in light of the challenges of social injustice, colonialism,
war, uneven development, pollution, resource depletion, loss of biodiversity, climate change, animal rights, and so on.

Yet even as geographers engage in ethical discourse, they still tend to speak about ethics as something external to the field itself, as if it were an extradisciplinary add-on. This is far from the case. Ethics has been part of the geographical tradition since the beginning; the moral turn is both an extension and a recovery of part of our intellectual heritage. For better or for worse, a moral voice has always been present in geographic inquiry.

These moral geographies are of many sorts. They include classical regionalizations of cultural diversity, critiques or justifications of imperialism, teleological explanations of the natural world, and the norms of social Darwinism that underwrote environmental determinism. The contemporary concern with explicitly theorizing and deploying ethical concepts in geography is a welcome addition, representing a more explicit and reflective disposition. As such, these deployments of ethical discourse represent a shift from implicit moral geographies to explicit geographical ethics.

Considered in this way, then, moral queries did not somehow infiltrate geography but have been there all along.

**Practical Reasoning**

One of the reasons for geography’s long connection with moral understanding is the field’s emphasis on context and contingency. Ethics has historically been a form of practical reasoning, which features context and contingency as central elements of causal explanation and moral justification. Practical reasoning differs markedly from the analytic reasoning that dominates modern moral philosophy. Modeled on the axiomatic sciences of mathematics and formal logic, analytic ethicists seek trans-geographical truth applied acontextually to the world. That is to say, they strive for universal axioms of conduct, derived without reference to the real-world experiences that occur over space and time. They then apply these abstractions to all people, places, and circumstances. The result, of course, is rigorously intended, if rigid, overinterpretations that are out of step with the world.

This is not the case, however, for the practical reasoning that is part of alternative traditions of ethics (e.g., casuistic, feminist, hermeneutic, theological). These alternatives thrive outside philosophy, as well as pose a challenge to analytic ethics within philosophy itself. Practical reasoning seeks to articulate situationally sensitive principles to guide us in moral and political deliberation. In this view, ethics is not a timeless and placeless body of truths but refers to the use of moral concepts as rules of thumb that help us understand how we ought to live. In this respect, geography is constitutive of ethics, generating conceptual and contextual insights that inform moral theory and method. Examples include ideas of space, place, and nature that have been and continue to serve as presuppositions to moral discourse.

**Ethics in the Internal and External Domains**

When it comes to the use of ethics in geography, there are two domains of ethical significance to consider. The first is the internal domain, that is, the methods of research and the production of knowledge. We often hear this domain explicitly referred to in terms of professional ethics, codes of conduct, or best practices. Ethics in the internal domain helps ensure the integrity and credibility of the field. While there are many ways of discussing this domain, it basically serves to uphold two moral values of science and scholarship—truth and trust. When speaking of truth, we are referring to matters such as the collection, analysis, interpretation, and communication of research. With respect to trust, we are thinking primarily about academic freedom, honesty, transparency, collegiality, and conflicts of interest. Along with upholding truth and trust as prime values, ethics also helps us define best practices for implementing these values in research. Common examples of best practices include the prohibition of plagiarism, falsification of data, and manipulation of research results, as well as guidelines on avoiding or disclosing conflicts of interest, prior restraint of knowledge, and self-censorship.

The second is the external domain, referring to the uses of geographical knowledge and the applications of its theories, methods, and associated technologies. We often hear this domain implicitly referred to when people speak about social
justice, environmental protection, sustainability, protecting local livelihoods, and so forth. Each of these phrases names a vital concern that embodies a substantial moral dimension. The reason for this external domain is that geography, for better or worse, has direct and indirect impacts on the health and well-being of people, animals, and the rest of nature. These impacts have consequences at a number of distinct if interconnected scales on individuals, populations, species, and communities, in natural and social systems, and in geographic space and historical time. Ethics helps elucidate the best uses of geographical knowledge by noting how research practices and knowledge products contribute to well-being in the world.

Although one may be tempted to classify the above domains into technical and critical modes, doing so unreflectively does an injustice to the diversity of intentions manifest in the work of geographers. While it may be true that, on balance, most scholarship in geographic information systems or qualitative research emphasizes the internal domain, there is important work in these areas that references the external domain. Some geographers attempt to balance the two in their research. And some emphasize one domain because they recognize its connection to the other. For example, the transparency of the research process in GIS is crucial if we are to use accurate mappings of nature and society for the greater good. And compliance with internal review boards in research with human subjects is a cumbersome but frequently necessary tool to protect the well-being of people participating in research. The domains should be considered mutually informing distinctions, not boxes into which we categorize (and perhaps dismiss) the full range and legitimate diversity of geographic research.

Challenges for Ethics in Geography

Having said all this, ethics still faces and poses significant problems in the discipline. Two of the more important are highlighted below.

The first of these problems is the empirical objection. The idea behind the empirical objection is that geography is a science committed to empirical research. Ethics is therefore not real geography as it is not “grounded,” “concrete,” “spatial,” “material,” or “factual.” This objection is not quite a rejection as it allows for a professional ethics enclosed within the internal domain noted above. It does, however, reject “speculative” excursions in so-called theoretical terrains, such as animal ethics, global ethics, and environmental justice. The error behind this objection is that it elides a discredited scientism and its empiricist vision of science with the tradition of geography as an interdisciplinary body of scholarship. This problem is compounded by an implicit facticity that makes invisible the causal relations between tangible and intangible phenomena. Both tangibles (e.g., the spatial patterns and extent of clear-cut forests) and intangibles (e.g., the values behind policy debates about the clear-cutting of forests) are real and thus empirical in any sensible definition of the term. To understand the intangible dimension, much less explain much of the tangible world, one must examine the moral causation that is partially constitutive of what and why humans do what they do.

The second problem is the challenge posed by animals and nature. After the debunking of social Darwinism and environmental determinism in the early 20th century, the field turned away from environmental matters for many decades. This has been redressed to some degree since the mid 1990s through technically focused environmental geographies (e.g., environmental geographic information systems [GIS]) and politically focused critical geographies (e.g., political ecology).

Yet from a moral perspective, these discourses remain stubbornly speciesist and anthropocentric. The reasons for this are many but are partially rooted in the value-free scientism that still stalks aspects of geography, as well as the social reductionism and moral relativism that characterize the social construction of nature thesis embraced by much of critical geography. To be sure, creative efforts to break out of these moral dead ends have emerged, primarily in animal geography and by dissident voices in nature-society relations. Their moral sensitivity to the well-being of people, animals, and nature is forging new insights that honor and extend the insights of technical and critical geographies while at the same time contesting the self-privileging moralities of human exceptionalism.

Overall then, geography is both the root and the fruit of moral understanding. As the root, its
situated knowledge is a necessary element of an ethics engaged with the real world. As the fruit, it is a conceptual space in which a more situationally sensitive ethics might thrive, even if (or because) it lends itself to a different model of practice than is normally pursued in philosophy. In this sense, geography shares with other disciplines (e.g., animal studies, environmental studies, and philosophy) the cotradition of ethics. As such, it also has a responsibility for developing geographical ethics that enrich our moral understanding of the world.

William S. Lynn

See also Animal Geographies; Critical Human Geography; Environmental Ethics; Environmental Racism; Existentialism and Geography; Humanistic Geography; Human Rights, Geography and; Inequality and Geography; Justice, Geography of; Marxism, Geography and; Phenomenology; Race and Racism; Social Justice; Spatial Inequality

Further Readings


Although there is no definitive definition, most scholars increasingly agree that ethnicity derives from both an internal sense of distinctiveness and an external perception of difference. The category exists to classify various groups of people based on specific social and cultural characteristics, with the most typical identifier being ancestry. Although some researchers continue to assert the emotional, hereditary, and primordial origins of ethnicity, increasingly there is agreement that ethnicity can also be the product of structural forces, social organization, and cultural representation. In other words, ethnicity is a social construction, where individuals are active agents in defining their ethnicity, and at the same time, the category must be negotiated within a reactive, shifting social environment.

Ethnicity is situational and dynamic, with individuals sustaining and asserting their ethnic identities in uneven and differential ways, depending on the social and political environment that surrounds them. So even though individuals may use the same ethnic label, they may construct their ethnicity based on the shifting notions and interpretations of their personal identities. At the same time, ethnicity is not a static concept that remains stable over time; instead, identities can be altered, manipulated, and transformed based on broader spatial, political, social, and economic dynamics. Ethnicity, then, is a creative and complex response to both individual and social forces. The formation and expression of ethnic identity come from both historical circumstances and individual negotiations to endow ethnicity and ethnic symbols with meaning.

Many individual characteristics are considered the building blocks of ethnic identity, including language, dialects, religious faith, literature, folklore, music, food preferences, social and political ties, traditions, values, and symbols, kinship, neighborhood, community links, and/or migratory status. Other external attributes are also thought to be significant in the construction of ethnic identities, including the role of governmental policies and social measures, racial discrimination, residential segregation, occupational concentration, and economic
isolation. Ultimately, each individual conceives, maintains, and makes manifest his or her ethnic identity in numerous ways: For some, the identity may be primordial and unproblematic; for others, ethnicity may be strategic and intentional. But all ethnic identities are transformative, and all ethnic identities have very real consequences in people’s everyday lives.

The Role of Race

No conception of ethnicity could explain much if it failed to reckon with the underlying issues of race. In contrast to ethnicity, the definition of which centers on cultural characteristics, race is taken to mean the common physical attributes of a given group, typically including skin color, hair texture, and facial features; however, racial differences are biologically almost meaningless, and race, like ethnicity, is also a social construction. All minorities are considered to have an “ethnic” identity, whereas those in the majority typically view themselves as “nonethnic.” But in most situations, this idea of not having an ethnic identity can only work for those in a privileged position. Everyone has an ethnic identity; the question is whether this identity is acknowledged. Nor are all ethnicities equal; for some, ethnicity might appear to be largely symbolic, costless, and voluntary; yet for many, there is no choice when it comes to defining their ethnicity— it is an identity, often conflated with race, that is imposed by the larger society. In other words, the dominant group can enjoy a considerable range of identities, suggesting that race allows it flexibility in choosing its ethnic identity, whereas the “visible” minorities have no such privilege: Even as they gain socioeconomic mobility, they are subject to stigmatization and marginalization based on their ethnic/racial identity.

Dominant groups play a critical role in the social construction and classification of ethnic groups as racialized minorities. Indeed, how the dominant group “sees” others is often more critical in the creation of ethnic/racial categories than how individuals identify themselves. This is primarily because the dominant group reserves for itself the luxury of dismissing ethnic/racial identities, while more oppressed groups are made daily aware of their ethnic/racial identities.

This ethnic/racial “labeling” becomes quite obvious when looking at how ethnic identity formation is intertwined with the process of migration. Immigrants frequently arrive in a new place where the dominant society exercises much power in defining difference. In an effort to not only survive but also prosper and build meaningful relationships, immigrants oftentimes depend on various ethnic attachments to create a sense of belonging. Typically, they become connected to those individuals with whom they share the most commonalities (in terms of locality, dialect, class, and religion, among other things). Yet because the host society relies on its own set of ethnic/racial categories, which generally focus on the country of origin (or even the broader world region), the diversity that characterizes individuals when they emigrate is disregarded for the categories that the host society uses to classify different groups. As a result of interactions with members of the dominant group and other minority groups in the host society, pan-ethnic labels (such as Indian and Mexican, or Asian and Hispanic) become concretized, and individuals of disparate origin oftentimes unite around these simulated categories despite their lack of resonance and meaning for individual immigrants.

Eventually, these distinct ethnic identities and group attachments should disappear as the descendants of immigrants take on the core culture and become assimilated. As members of the second and beyond generations begin interacting with members of the dominant group, particularly on the intimate levels of friendship and family formation, ethnic differences should become less and less significant. Ethnicity should remain important in only the most superficial ways after assimilation occurs, as individuals choose symbols (such as ethnic foods and festivals) that require little effort to maintain and that do not interfere with other social and economic aspects of their lives. This symbolic ethnicity emerges as a nostalgic connection to the culture of the “old” country, whereby individuals carefully and continually (re)choose those ethnic traditions and ideals that serve them in varying contexts. While this idea appears to be valid with regard to ethnic groups that “blend in” with the host society, the reality is that groups that appear to be different (particularly
in terms of physical attributes, religion, and class) have little choice in deciding their ethnic/racial identities. Some “visible” minorities are even forced to appropriate these involuntary labels to use collectively in group actions since society is organized and ruled along ethnic/racial lines.

**The Role of Space**

Recent work argues that the study of ethnic identities cannot be understood without explicating the relationships between ethnicity, space, and place. This geographic research takes shape in many forms, although only two areas are discussed here: (1) mapping the spatial distributions of ethnic groups and categorizing various ethnic landscapes and (2) identifying how the regulation of space reflects and reinforces ethnic/racial identities.

Many geographers pay special attention to the spatial distributions of various ethnic groups. An entire terminology has emerged to talk about the various forms that ethnic communities take on, including ethnic ghettos, enclaves, and citadels. In this work, typically, the unifying force of ethnic attachment in creating and re-creating ethnic communities is highlighted. Ethnic ghettos, ethnic enclaves, and ethnic citadels are distinguished by their homogeneity, density, and old-country “flavor.” The emotional links of ethnicity bind residents together and refortify the continuation of ethnic communities. Each distinctive ethnic group creates and occupies a place that becomes representative of that group. These places develop into vibrant ethnic centers that not only concentrate and display a distinct group’s way of life but are also supported by specialized institutions and commercial activities.

Often, the spatial distribution of ethnic groups is expected to indicate broader patterns and processes of settlement and adaptation. At the heart of this research is the notion that residential change and social mobility are connected and interlinked. Here, spatial dispersion (and the disappearance of concentrated ethnic communities) is indicative of the weakening of social and structural barriers, thus indicating not only spatial but also social and structural assimilation into the dominant society. This geographic research illustrates that ethnic distributions are not necessarily stable and that spatial patterns and levels of segregation among and between ethnic groups somehow measure the degree to which groups have become integrated.

Because the primordial and emotional connections of ethnicity tend to be overemphasized in this research, some geographers are revisiting these mappings. Recent critiques and theorizations have exhibited a greater sensitivity to the role of structural forces in reinforcing ethnicity and ethnic landscapes. There is a particular focus on the role of space in creating and re-creating ethnic identities. Scholars articulate how various groups experience places as inclusive or exclusive of them and others and how the regulation of space reflects and reinforces the identities, privileges, and interests of some groups over others.

New questions centered around issues of ethnicity focus on (1) the forms of racism and racial privilege and oppression that operate through geographical patterns, processes, and ideas and (2) the ways in which ethnic/racial hierarchies become struggles over space within society. Thus, ethnic ghettos and ethnic enclaves (along with ethnic citadels, which are the ethnic landscapes of the dominant culture) are considered the sites and mediums of privilege, power, and prestige, as well as oppression, minoritization, and disparity. Ethnic groups positioned in the most impoverished and outcast spaces, as well as those living in the wealthiest, most exclusive places, are spatial reminders of broader ethnic/racial divides (and socioeconomic hierarchies as well). Ethnic/racial categories are inhabited and naturalized, both materially and ideologically, through these concrete spaces, primarily as a result of systematic institutional practice and exclusionary spatial codes.

Certain ethnic groups become racially categorized during the time when the neighborhoods they live in become spatially segregated communities. But the spaces of the dominant group, which are usually taken for granted as invisible, ordinary, and mundane, are also spatially segregated landscapes, which in turn reinforce the ethnic/racial category of privilege as one that should be considered normal, taken for granted, and unmarked. This process of differentiation occurs because the members of the dominant group
cannot (or refuse to) acknowledge how they themselves embody particular ethnic/racial categories. By maintaining invisibility, the dominant group remains racially undefined, and social hierarchies remain rigidly reinforced. Thus, the ways in which ethnic spaces in turn both produce and reflect specific racial ideologies become critical. Even though ethnic/racial categories and geographies change with time, they continually matter because groups are relegated to different points along the spatial hierarchy, which in turn reinforces disparity.

### Conclusion

It is important to point out that despite its wide usage, ethnicity as an analytical concept in academia still lacks a universal definition. The term *ethnicity* is used with varying meanings in different contexts, and these understandings change over time and across space. Moreover, ethnicity is becoming increasingly complex as the crossnational and transnational movement and connections of people, goods, information, and financial capital become more prevalent. Certainly, ethnicity affects societies deeply and in multiple ways that are not always easily identified or well understood. Identifying the many influences of this categorization scheme (especially the role of race and space), as well as its transformative and dynamic nature, can begin to help untangle the influence of ethnicity in everyday life.

*Emily Skop and Wei Li*

---

**See also** Environmental Racism; Ethnicity and Nature; Ethnic Segregation; Ethnocentrism; Identity, Geography and; Immigration; Nation; Nationalism; Orientalism; Other/Otherness; Race and Empire; Race and Racism; Racial Segregation; Segregation and Geography; Urban Underclass; Whiteness

---

**Further Readings**


---

**Ethnicity and Nature**

The relationship between ethnicity and nature has often focused on a notion of essentialized identities. People and place are linked, in this view, not as mutually constitutive social and cultural constructs but rather as immutable and static objects and subjectivities. Ethnicity has often been viewed, therefore, as a way of expressing assumed relationships between particular communities and specific regional landscapes and resources or through the cultural and economic practices of a given population. Such linkages have sometimes been normalized and reified to the point that the particular racial, national, or ethnic heritage of a group begins to define a landscape itself or vice versa. We begin to talk therefore of an English countryside, a Canadian prairie landscape, or a Rajasthani desert. Similarly, we may begin to describe particular communities as forest dwellers, hill peoples, or desert nomads, privileging a view that echoes with the longdiscredited environmentally deterministic theory of human development. It is important therefore to interrogate some of the key discourses that...
have linked ethnicity and nature as an expression of essentialized identity. There are three particularly central themes in which this relationship is expressed: (1) nationalism and nature, (2) the noble savage, and (3) environmental racism and environmental justice, which this entry addresses in turn.

Nationalism and Nature

Despite the criticisms that have been levied against environmental determinism for several decades, the power of the assumptions embedded in the relationship between ethnicity and nature remains enduring. Modern nationalist movements, for example, from those of 19th-century Europe to the anticolonial struggles worldwide in the 20th century, have drawn consistently on the idea of inherent and essential connections between people and place as a way of imagining new political communities. We often find in Romanticism, the cultural and intellectual movement that undergirds many of these emerging nationalist visions, an explicit link between an ethnic and a physical nature—pastoral landscapes and an English character or perhaps dark forests and an essentially martial German subject. Similarly, the anticcolonial nationalism of the past two centuries has often drawn on the idea that an indigenous population is somehow more legitimately connected to the local landscape than is a “foreign” ruler. Such claims have often been made not only on the basis of justice and equity but also on the assumed authenticity of one group rather than another, an authenticity predicated once again on an essentialized notion of ethnicity and nature. The British, in this view, were not welcome in India, the French in Algeria, or the Dutch in Indonesia, not only because of their repressive behavior as imperial masters but also because the stain of their unfamiliarity with the natural landscape of each place could not be erased through colonialism. The British retreat to hill stations in the summer, the French attempt to re-create metropoles, and the Dutch rejection of the traditional longhouses were all therefore seen as an inability to connect to local landscapes.

In the contemporary period, ethnicity remains a potent fount for articulating “authentic” connections between specific people and places.

Indeed, many of the major political conflicts that have arisen throughout the 20th century have been characterized by struggles to define the “natural ethnicity” of a region. Israel and Palestine, for example, are seen from this perspective as being quintessentially Jewish and Arab, with these ethnicities defining the land and vice versa. So-called ethnic cleansing has also been used in conflicts, from the former Yugoslavia to Rwanda, and many others besides, to remove apparently offending bodies and communities from lands to which it is claimed they do not belong. An essentialized understanding of ethnicity and nature lies once again at the heart of such political categorizations.

The Noble Savage

Romanticism not only celebrated the notion of specific landscapes as touchstones of nationalist identity but also venerated and elevated an earlier cultural archetype of an uncivilized and unspoiled individual or culture for its supposed closeness to and familiarity with nature. In the context of ethnicity, there are two distinct effects of the noble-savage characterization. On the one hand, the archetype has been used to justify paternalism, racism, and colonial subjugation by suggesting that certain ethnic identities are unable to control and manage nature effectively. On the other hand, the notion of the noble savage has at times become a fantasy of purity through which so-called civilized or modernized groups and individuals may lay claim to an unspoiled and natural past. Both views are demeaning, reductive, and pernicious, yet they have had distinct, if related, consequences.

The paternalistic view of the noble savage posits the idea that particular ethnic identities have what amounts to an unsophisticated and immature understanding of nature. This view does not necessarily suggest that these identities are not connected to or situated within nature; paradoxically, noble savages are often glossed with the sheen of a positive and benign association with natural phenomena and space (see below). But the fundamental characteristic that defines these particular understandings of ethnicity and nature is one of inadequacy: an inability to manage resources, to understand—in a scientific, rational
manner—the workings of the natural world, and to preserve and maintain an ecology in a sustainable fashion. Indigenous and tribal communities have suffered in particular because of this characterization and were often labeled as primitive, backward, and ignorant “others.” From the perspective of (self-described) advanced societies, such cultures are often deemed either incapable or unworthy of not only the responsibility of self-rule but also the stewardship of the natural environment in which they are situated. The noble savage in this case may live in nature but is to some degree ignorant of it and its proper use. Throughout the period of European imperialism, but especially during the late 19th century, this particular view of ethnicity and nature helped justify the expropriation of lands and resources from indigenous communities across the world in the name not only of conquest but also of environmental stewardship and the scientific management of natural resources.

Although the ideology of the “white man’s burden” may have faded in the postcolonial period, some critics have argued that primitivist understandings of ethnicity and nature continue to play a substantial role in the contemporary lives of many indigenous, tribal, and marginalized populations across the world. Indigenous communities have consistently found themselves displaced, for example, from traditional lands and livelihoods across the globe. This process has occurred as a result of the projects of nation-states, the profit motives of extractive industries, and the protectionist politics of conservation efforts. Whether because of the building of dams, the logging of forests, or the creation of parks, the same logic that links “primitive” tribal natures to their inability to manage their own environments underlies these displacements. If noble savages are present within so-called natural landscapes, they are often acceptable only as museum pieces or cultural artifacts rather than as active participants. Indigenous communities therefore may be transformed into particular attractions or performers in tourist or ecotourist destinations.

A more celebratory representation of the noble savage presents the reverse view to that of the ignorant and primitive Other, although it arises out of the same essentialized notion of race and ethnicity. In this view, indigenous and tribal identities are exoticized and fetishized precisely for their apparently deeper and more authentic connection to the natural world. Such Orientalized understandings have become a prevalent lens through which some Native American/First Nations identities have been framed based on an argument that these particular groups may be better stewards of the Earth because of the closeness that their ethnicity provides to nature.

Similarly, ideas regarding Eastern spiritualism—linked variously to traditions in India, China, Japan, or the Middle East—have been a significant source of New Age beliefs and in turn constitute an important influence on some aspects of the modern North American environmental movement. Relying once again on an essentialized view of ethnicity and nature, the positive representation of the noble savage—as both individual and community—has served as a model for environmental activism and the articulation of an alternative ecological vision. Indigenous communities are themselves not ignorant of such reductive views of themselves, in both the positive and negative portrayals as pure or ignorant. Indeed, some groups have engaged in what might be deemed “strategic essentialism,” playing off the stereotypical views to gain necessary support for the social and ecological justice struggles in which they are engaged. The longer-term costs of the accommodation of the noble savage representation by such subjects is, however, difficult to gauge.

**Environmental Racism and Environmental Justice**

In the same way that the noble-savage characterization has two distinct but related faces, the relationship between ethnicity and nature has a dichotomous form when we look at the issues of environmental racism and environmental justice. On the one hand, the issue of ethnicity is clearly visible when we look at the notion of environmental injustice as it has been popularized over the late 20th and early 21st centuries. Many activists and scholars have alleged that it is on the basis of ethnicity and discrimination in large measure that polluting industries and hazardous materials have been placed within racialized and marginalized neighborhoods, regions,
and countries. Similarly, it is on the basis of assumed linkages between essentialized identities and the natural world that certain groups have been pressed into service through specific forms of labor and exploitation, whether in sweatshops, extractive industries, or fields and factories. On the other hand, a demand for environmental justice has been a rallying cry in recent years for a range of social movements worldwide, many of which base their campaigns in part on the belief that particular ethnic identities—especially marginalized and indigenous communities in the so-called Third World—are especially under environmental threat. Both parts of this equation continue to reify the linkage between ethnicity and nature and similarly continue to have deleterious effects, especially in terms of global environmental politics and change.

Ironically, while the campaigns of many prominent environmental organizations have privileged indigenous struggles across the world, within North America the environmental movement as a whole has often been criticized as being a primarily white, affluent, and urban phenomenon that discounts or denigrates the possible contributions of minority groups. Such populations—including recent immigrants, poorer or working-class communities of color, and historically disadvantaged groups—are often characterized as having less knowledge or interest in environmental issues. At the same time, the increasing internationalization of certain environmental problems—such as global climate change, species endangerment, deforestation, and habitat destruction—has led to the search for transnational solutions, many of which often lead to the disenfranchisement of indigenous and otherwise ethically identified groups through a process that some observers have described as green globalism or green neocolonialism.

Pablo Bose

**Further Readings**


**ETHNIC SEGREGATION**

Ethnic segregation is the enforced or voluntary residential separation of two or more groups on the basis of cultural identity. Usually understood at the local scale among neighborhoods within a city, ethnic residential segregation also occurs at other geographic scales. Regionally, Native Americans have been forced onto reservations encompassing many U.S. counties, and many blacks in South Africa were relocated to several “homelands” during apartheid. The various aspects of ethnic segregation include what causes segregation, how segregation can be measured, and what interpretations can be made from segregated spaces. Overall, the separation of ethnic groups into distinct geographic spaces due to coercion or possible discrimination may illustrate the inequalities found in societies.

**Causes of Ethnic Segregation**

There are several causes of ethnic residential segregation: housing discrimination, socioeconomic
differences between ethnic groups, ethnic group preferences, and forced ethnic confinement. Housing discrimination, in both public and private sectors, has been proposed to cause ethnically segregated living spaces in American cities. In the past, the U.S. federal government attached provisions to publicly offered home loans that restricted the housing options of various groups of people. Local governments, via their zoning laws, tended to exclude certain ethnic groups from residing in the suburbs. In the private sector (i.e., banks, realtors), limitations to housing for particular ethnic groups were facilitated by redlining and blockbusting. Such formal restrictions are now illegal, but informal practices of housing discrimination may still exist today.

Another prospective cause for ethnic residential segregation deals with differences in socioeconomic status (SES). The basis of the SES argument is that residential segregation is related to class differences rather than to ethnic differences. Here, class issues underlie housing outcomes in determining where people live in a city. If economics were the only factor in determining where people choose to live, then the poor would live in the areas with older housing and the rich in the areas with newer housing, or in the inner city and the suburbs, respectively.

Ethnic group preferences for living in segregated areas are another hypothesized cause of residential segregation. Under this interpretation, ethnic groups differ in the type of neighborhoods they perceive as desirable. Ethnic residential segregation may signify social solidarity, if ethnic members choose to live near each other, or social avoidance, if people prefer not to live near a certain ethnic group. Self-segregation can be a method of creating defensible spaces that maintain “peace” between opposing groups, as the example of Catholics and Protestants in Belfast, Northern Ireland, indicates. The last prospective cause of ethnic segregation deals with the forced confinement of an ethnic group to a particular area. The first “ghetto” in Venice, Italy, was created for Jews by forcing them into a neighborhood that was separate from the residential areas of other Venetians. In terms of regional ethnic segregation, Native Americans and many Japanese Americans during World War II were relocated and residentially separated from other Americans.

### Measuring Ethnic Segregation

Determining the levels of residential segregation is important in monitoring whether cities are becoming more or less segregated over time. There are several segregation indexes that have been disputed over by social scientists, but the most used ethnic segregation index is the dissimilarity index (DI). The DI measures the differential distribution between two population groups in an area and can be calculated using the following formula:

\[
\sum_{i=1}^{n} \left| \left( x_i - y_i \right) \right| \cdot \frac{1}{2}.
\]

In this formula, \( x_i \) is the percentage of the total \( x \)-group population in census tract \( i \) and \( y_i \) is the percentage of the total \( y \)-group population in census tract \( i \). Here, the sum of the absolute values of the differences in percentages between group \( x \) and group \( y \) values in census tract \( i \) is divided by 2. The DI ranges from 0, indicating no segregation between the \( x \) and \( y \) groups, to 100, indicating complete segregation between the \( x \) and \( y \) groups. Calculating and mapping this segregation measure between African Americans and whites for American metropolitan areas (Figure 1) illustrates that cities in the Midwest and the northeast have higher segregation levels (with DI greater than 60) than cities in the west.

The drawback of the DI, for quantitative research using enumeration units, is the modifiable areal unit problem. Geostatistical analysis is sensitive to the sizes of the enumeration units used in its calculations, since spatial data can be enumerated for multiple spatial scales (e.g., county level to tract level to block level). Depending on what enumeration unit size or shape is used to measure segregation, different segregation scores can be computed.

### Meaning of Ethnic Segregation

There are positive and negative implications of living in ethnically concentrated areas that are
conceptualized in the dichotomy of the ethnic enclave/ghetto paradigm. In other words, there are “good” and “bad” forms of segregation; living in the former provides support for group members, whereas living in the latter hinders future life chances for group members. Ethnic enclaves, or the “good” segregation, have been represented historically in the form of urban places such as Little Italy, Chinatown, and the like, which have provided community support for people with similar ethnic backgrounds. In essence, the ethnic enclave was a “home away from home” for immigrants that made transitioning into another society easier.

In general, the stigmatized ghetto was characterized as a place that had poorer quality-of-life outcomes (e.g., poor education, high poverty and unemployment rates) and where ethnic groups without choices lived. The ghetto was associated with high crime rates, areas of social vices, and places that housed the most downtrodden in American society. Living in the ghetto was detrimental for its residents, in that trying to make it by moving out was difficult due to the lack of social capital in these neighborhoods. Ghettos were seen to remain on the urban landscape over time since it was too difficult to change the downward spiral of social ills.

Ethnic segregation between African Americans (Figure 2) and whites (Figure 3) in Omaha, Nebraska, can be related to the location of poverty (Figure 4). Areas of “North Omaha” have a high percentage of African Americans and high rates of poverty, while areas of “West Omaha” have a high percentage of whites and low rates of poverty. This spatial association

**Figure 1** Levels of segregation between African Americans and whites for each American metropolitan area in 2000

**Sources:** Map created by author, based on data from U.S. Census Bureau, www.census.gov, and Lewis Mumford Center, http://mumford.albany.edu/census.
Figure 2  African Americans in Omaha (2000)

Figure 3  Whites in Omaha (2000)
between ethnicity and one quality-of—life measure indicates that segregation may benefit some ethnic groups over others. However, on the positive side of African American segregation in Omaha, several ethnic businesses and organizations have found a place to flourish in North Omaha. Also, community cohesiveness in ethnically segregated spaces may be high, even with the poor socioeconomic characteristics. Overall, living in ethnically segregated areas can have both advantageous and disadvantageous outcomes for its residents.

*Kenneth French*

*See also Chicago School; Class, Geography and; Ethnicity; Ghetto; Housing and Housing Markets; Identity, Geography and; Neighborhood; Poverty; Race and Racism; Racial Segregation; Segregation and Geography; Urban Spatial Structure; Urban Underclass*

**Figure 4** Poverty in Omaha (2000)


Further Readings

ETHNOCENTRISM

Ethnocentrism means that a group of people views the world from its own distinct set of cultural values, or ways of living and thinking. This bias typically leads to feelings of superiority in its relationships with other groups. Ethnocentrism varies widely in type and degree among cultures.

Ethnocentrism makes appreciating cultural diversity difficult if not unattainable. Indeed, most people do not understand that they do not understand other people’s cultural values. People may try to be genuinely empathetic to those with different cultural values, but even geographers, anthropologists, and sociologists fail in ridding themselves of subconscious predilections.

Bigotry and racism can emanate from ethnocentrism. Extreme ethnocentrism can cause ethnocide or genocide. Ethnocide is the killing, either unintentionally or intentionally, of a group’s culture by another group. For example, the construction of the Trans-Alaska Pipeline in the 1970s inadvertently contributed to the acceleration of ethnocide of the native cultures. The ethnocentric “white man’s burden” justified European colonialism throughout Africa, subjugating its cultures and enslaving its peoples. Ethnocentrism sparked the German Nazis to kill 6 million Jews during the Holocaust in the 20th century. Ethnic cleansing is a process by which a dominant group of people expels a minority ethnic group from a particular place. Ethnic cleansing drove Bosnians from their ancestral homes in recent warfare in Southern Europe.

Ethnocentrism inscribes the landscape. China, “the Middle Kingdom,” considered itself for centuries as the center of its universe and people located farther from its cultural hearth, for instance, in Inner Mongolia and Outer Mongolia, as increasingly more barbaric. Europeans stamped their “superior” cultural traits onto landscapes around the globe as they conquered and controlled numerous native peoples. Extractive processes, agricultural fields, urban structures, and religious icons changed indigenous homelands indelibly. These cultural conquests set the stage for a world system of core-periphery relations and hegemonic actors dominating the marginalized aborigines. Ethnocentrism plundered the resources of the peripheral places and decimated native cultures and their natural environments.

Ethnocentrism was central to the idea of “manifest destiny” in U.S. historical geography. In their march westward to the Pacific Ocean, Anglo Americans vanquished the Native American, French, and Spanish ways of living. “The Five Civilized Tribes” were not “civilized” enough to protect their sacred lands and prevent their children from dying along the “Trail of Tears.” Anglo-dominated Louisiana legislatures promulgated English-only laws for all government proceedings and school instruction to eliminate the remnants of French culture. Under the Treaty of Guadalupe-Hidalgo, the United States annexed half of Mexico and imposed new mores, especially language and law, on the inhabitants captured in the new country.

Ethnocentrism, then, is an extraordinarily powerful cultural concept that not only leads to misunderstandings among people but is also responsible for devastating natural environments, destroying nations, and killing millions of people.

Lawrence E. Estaville

See also Discourse and Geography; Ethnicity; Eurocentrism; Identity, Geography and; Nation; Nationalism; Orientalism; Other/Otherness; Positionality; Race and Racism; Social Darwinism; Situated Knowledge; Whiteness

Further Readings


Usage of the terms Eurocentrism, Eurocentrist, Eurocentric, and Eurosceptic within the social sciences and geography has increased since the 1990s, within the West and in areas of the world that were formally part of the overseas colonial empires of European powers. This entry reviews the definitions, origins, history, and geographical implications of Eurocentrism and its subdivisions. It then describes the counter-Eurocentrism movement.

Definitions of Eurocentrism

There are three interrelated definitions of Eurocentrism. However, it is the third and last definition that reflects by far the most common usage of the term.

Definition 1. In Britain, a Eurocentric or Eurocentrist is a person who supports European integration or the concept or workings of the European Union. The opposite of a Eurocentric is a Eurosceptic.

Definition 2. The term Eurocentrism is sometimes used to convey the perspective of North Americans (the United States and Canada) and may occasionally be defined as the viewpoint of North Americans of European origin. Some scholars have begun using the specific term EuroAmericocentrism or simply Americocentrism in this context.

Definition 3. Eurocentrism is a view of the world from a perspective that places Europe and its cultures, concerns, and sociopolitical attitudes at the center of the world. This may be done consciously or subconsciously, but it is always to the detriment of non-Europeans. It can be argued that true Eurocentrism has existed only since World War II and the development of first the European Economic Community (EEC) in Western Europe, followed by the European Union (EU), which has been expanded by the accession of the former Eastern European states. The same argument suggests that prior to World War II there were only national policies and national centrisms.

Since the 16th century there has been a growing socioeconomic and political hegemony of European nations (principally, in chronological order, Spain, Portugal, the Netherlands, Britain, France, Italy, and Germany) over large areas of the world. This status has led to Western European thought permeating many disciplines.

Perhaps the most common form of Eurocentrism is the division of the Eurasian landmass into the “continents” of Europe and Asia at the Ural Mountains—although there is no objective basis for this separation. Indeed, this long-accepted division was the invention of Vasilli Tatischev, an 18th-century geographer for Peter the Great, who used it as part of the Russian elite’s attempt to differentiate Europe from Asia as meaningful entities. By this definition, Europe becomes the only “continent” or major world cultural region that is not separated by water from other so-called continents (although what exactly constitutes a continent is arguable).

In the same vein, within cartography, the Mercator projection distorts the equatorial regions and exaggerates the sizes of polar regions. This projection gives the impression that European nations are larger in areal size (and therefore more important) than countries located nearer the equator.

Toponyms or place names also show a particularly strong Eurocentric resilience. For example, the use of the term Middle East (French: Moyen Orient) is acceptable only if one assumes that the user of the term is located to the west of this region. Such a term makes the Far East (French: L’extrême-Orient) another Eurocentric toponym; a region is in the Far East only if one is located in the “West.” In contrast, Arab geography divides Western Asia into distinct geographical and cultural areas and has never viewed the entire region as a culturally homogeneous region. Within political geography, the application of the terms the West or the Western world is Eurocentric in that they imply a certain political and cultural territorial construct rather than the physical reality. The chronologically and spatially fluid terms the West and the Western world are often perceived as referring to North America, Europe, Australia, New Zealand, and Japan.

A commonly noted example of a Eurocentric theory is Max Weber’s concept of the Protestant work ethic. Weber held that industrial capitalism
began in Northwestern Europe because its values, stressing hard work and delayed gratification, were inherently superior to those of other cultures such as Islam, Hinduism, and Buddhism. Weber’s ideas became widely influential in later modernization discourses.

### Origins of Eurocentrism

Eurocentrism has its roots in the European Renaissance, with its celebration of Greek and Roman civilization, often with the omission or denigration of African (e.g., Egyptian, Carthaginian) or Asian (e.g., Sumerian, Assyrian, Persian, Indian, or Chinese) cultures. Eurocentrism has increased since the colonial voyages of discovery and the subsequent period of European colonial and empirical expansion across the world.

In the Iberian peninsula, Eurocentrism reached its zenith during the 17th and early 18th centuries, where it was part of the doctrines of religious and racial superiority (Portuguese, Raças Infec\-tas/Limpeza and Pureza de Sangue; English, Contaminated Races and Purity of Blood). In Northern Europe, similar expressions of Eurocentrism reached a high point in the 19th century, where the notion was seamlessly meshed with theories of European racial superiority, such as social Darwinism, as well as the concepts of race-based slavery and technological, political, and economic domination of non-European peoples. In some cases, Eurocentrism was manifested in illusions of grandeur and exoticism (as in the notion of the noble savage), but by the early 20th century, Eurocentrism had become part of the machinery of empire and European economic and cultural domination over large parts of world. Eurocentrism reflects the relative strength of European economic and sociopolitical power. Europe has always been composed of heterogeneous sociopolitical and economic state entities. Therefore, there is a spatial and chronological fluidity to Eurocentrism. The concept of Eurocentrism has traditionally applied to Western Europe and is rarely inclusive of Russia.

### Subdivisions of Eurocentrism

The origin of Eurocentrism lies in the concept that European classical (specifically Hellenic and Roman) civilization developed in a unique and culturally isolated manner. The contra-Eurocentrist movement has contested this notion for some time. The archaeologist Gordon Childe first challenged Eurocentric theory. In the 1930s, Henri Sée and H. M. Robertson opposed aspects of Max Weber’s theory that postulated that only Protestant countries could achieve economic progress.

The end of World War II and the subsequent collapse of European empires marked the start of another political, counter-Eurocentric, movement, including the establishment of the United Nations, with its five permanent members (France, the United States, the United Kingdom, the USSR, and China), in which the European nations were a minority.

Eurocentrism was particularly challenged in the Francophone world by the concept and movement of Negritude, which began in the 1930s and was advocated by the Martinican poet Aimé Césaire, the future president of Senegal Léopold Sédar Senghor, and the Guianan political poet Léon Damas. The Negritude movement, which affirmed the value of African culture and emphasized the shared heritage of blacks through the African Diaspora, voiced a perspective contrary to that of the European colonizing powers.

In the 1970s, a similar movement, termed *subaltern studies*, arose in South Asian historiography. First, *subaltern* was the term used for minor functionaries of the British colonial regime in India, including Indian, Anglo-Indian, and English officials of a low rank in the British military forces. Second, and more important, it is a term often associated with the work of the Italian Marxist political philosopher Antonio Gramsci, who used it to refer to the working class, peasants, and proletariat, who were excluded from, and who stood in opposition to, the dominant historical bloc of a given social formation. More recently, subalternism has given way to the term *Indocentrism*, perhaps underlying the fact that national centrisms are reflections of the increasing socioeconomic power within the regional or world order (Table 1). After the rise in postcoloniality, Eurocentrism has been challenged at its core historical basis by research into the African and Asian contributions to classical European civilization by Martin Bernal.
Through the 1980s in Europe, especially in France and Britain, Afrocentrism, multiculturalism, and so-called political correctness have challenged Eurocentrism. Much of the force behind these movements has come from the African, Caribbean, and Asian diaspora in Europe.

In geography, Eurocentrism has been strongly criticized by scholars such as James Blaut, who dismiss the “tunnel vision of history,” which he claims asserts privileges for European culture over others. The analysis of Orientalism, initiated by Edward Said, has likewise led to the rise of discourses in which the West is seen as a historically specific configuration of power relations rather than some inherently superior entity. Through the 1990s and the early 21st century, there has been an emphasis on viewing the diaspora and the accompanying culture contacts or cultural (and genetic) interchanges as creating fluid hybrid or “Creole” cultures and peoples of multiple heritage. This notion of hybridity is increasingly perceived as non-Eurocentric, equitable, and diverse and therefore more acceptable as a basis for political and social cohesion.

*Clifford Pereira*

<table>
<thead>
<tr>
<th>Regional Centrism</th>
<th>Subregional Centrism</th>
<th>National Centrism</th>
<th>Provincial Centrism</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eurocentrism</td>
<td>Iberocentrism (Iberian Peninsula)</td>
<td>Anglocentrism (United Kingdom)</td>
<td>Anglocentrism (sometimes only applied to England/London)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Francocentrism (France)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Lusocentrism (Portugal)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hispanocentrism (Spain)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Germanocentrism (Germany)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Italocentrism (Italy)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hispanocentrism (sometimes applied only to Madrid)</td>
<td></td>
</tr>
</tbody>
</table>

**Table 1** Structure of Eurocentrism

Source: Author.

Further Readings


The Euromarket refers to a range of financial markets trading in currencies and bonds. Although there are significant differences between these markets, their common feature is that they trade in products denominated in “national” currencies outside the jurisdiction of the relevant state. Euromarkets are geographically significant because they have altered and extended the meaning of economic space. This entry reviews the origins and evolution of the Euromarkets, their role in the global financial system, and their
implications for the practice and study of economic geography.

There is no precise date for the creation of the Euromarkets, but most commentators agree that they came into existence in 1957, when the Eurodollar market was created in London. The Eurodollar market emerged at this time because of a combination of the U.S. dollar's dominance as an international currency, the dynamics of the Cold War, and the strictly regulated nature of U.S. banks. The Marshall Plan had pumped large amounts of U.S. capital into Europe to fund reconstruction, Japan was occupied by U.S. forces and was also being rebuilt, U.S. military spending was rising rapidly, and international trade was growing. Despite this, the U.S. banking system was both bureaucratic and restrictive, placing limits on the profitability of U.S.-based reserves and on the liquidity and flexibility of European banks. Many European banks—particularly in London—wanted an alternative. Further pressure on the system came from Soviet bloc countries, particularly the USSR, with dollar reserves derived from oil production that they did not want to expose to sequestration by a hostile U.S. government.

The Eurodollar market resolved these various problems by allowing banks to use their dollar reserves to make loans to each other and to clear dollar balances entirely outside the U.S. banking system. Because of a peculiarity of English law whereby if money is not formally allocated to a particular legal jurisdiction it is effectively nowhere, the Eurodollar market created an unregulated and untaxed legal “space” within which U.S. currency could be freely traded. Although originally confined to European banks, as the Eurodollar market grew, it quickly attracted U.S. banks wanting to take advantage of the new market’s ability to circumvent the Federal Reserve Board. Eurodollars were soon joined by many other Eurocurrencies and by the Eurobond market. The prefix “Euro” no longer refers either to Europe or to the currency euro but to any financial product denominated in any currency outside its formal jurisdiction.

Although it was probably not intended by those involved in its creation, the Euromarkets have fundamentally altered the economic geographies of the 20th and 21st centuries. The markets have grown exponentially, stimulated by the rise of the petrodollar in the 1970s and the deregulatory policies of many Western governments in the 1980s and 1990s. The Euromarkets are often cited as key drivers of the process of economic “globalization.” Estimates of the values of the markets vary because of their unregulated and volatile nature, but it has been calculated that as much as $3 trillion to $5 trillion circulates in the markets daily. The Euromarkets have become key components in the offshore economy, many Euromarket products and reserves being held by companies based in tax havens such as the Cayman Islands, Jersey, or Switzerland. The Euromarkets have also become highly automated, using complex market algorithms and information and communications technologies to reduce transaction times to milliseconds.

For geographers, the Euromarkets are of importance since they operate within an entirely virtual monetary space above and beyond any conventional economic or legal territory. This greatly adds to the complexity of economic geographies since they can no longer be conceived of in simple spatial terms. The majority of the “Russian” economy, for example, which has benefited from many billions of U.S. dollars in aid and oil revenues since the collapse of the Soviet Union, largely resides offshore in the Euromarkets and the world’s tax havens. Although the Euromarkets are a relatively new phenomenon, the spaces within which they operate are not. As a number of commentators have argued, the placelessness of the Euromarkets was also a feature of the world economy prior to World War I, when capital was able to move across borders without let or hindrance. As such, the idea of a territorially defined sovereign nation-state managing a bounded economic space appears as historically and geographically anomalous.

This does not mean that the state’s role in the economy has been eradicated by the Euromarkets. The virtual legal space these markets occupy is defined in distinction from sovereign state spaces. As such, the Euromarkets need state jurisdictions to create and maintain “national” currencies, to guarantee property title, and to participate in the market for residency and citizenship (personal and corporate), from which much of their profitability is derived. For their
part, although states lose potential tax revenues to the unregulated realm of the Euromarket, they are also significant market participants, holding much of their “national” wealth in the form of extraterritorial financial instruments. Although recent problems in the world financial system have led to calls for greater regulation of the Euromarkets (e.g., the Tobin Tax), these placeless economic spaces remain centrally important to the geographies of the world economy.

Angus Cameron

See also Circuits of Capital; Economic Geography; European Union; Finance, Geography of; Globalization; Offshore Finance; Money, Geographies of; Telecommunications and Geography

Further Readings


**EUROPEAN GREEN MOVEMENTS**

The European green movements are the individual diverse movements and struggles that emerged in the post–World War II period through explicitly politicizing aspects of the environment. These movements have taken markedly different forms in different places and times and have been organized around multiple grievances. One of the striking aspects of the green movements is their political heterogeneity, and many green activists have explicitly disrupted notions of left and right. Members of the German Green Party (*Die Grünen*) famously entered the Bundestag in 1983 with the slogan “Neither left nor right—but straight ahead.” The European green movement has made important interventions in the style of doing politics and what counts as politics. It comprises a diverse set of organizational and political forms including direct action, parliamentary parties, nongovernmental organizations, and local environmental campaigns and has had a major impact in making the environment and biopolitics central political issues of our times.

These movements emerged from and have developed a set of plural and contested trajectories. These include the (often rather conservative) conservation groups as well as the more explicitly countercultural politics associated with the New Left. The relations between the European Green Movement and the New Left demand significant attention. Thus New Left activists of the 1960s and 1970s often translated their political activity into emerging social movements such as the German Green Movement and played a key role in articulating green politics away from an authoritarian, conservative agenda. Similarly, the currents of radical politics established in 1968 in Paris had underground afterlives in the work of activists such as Jose Bové, who shaped Left articulations of environmental politics in rural France.

New Left currents also shaped the emergence of environmentalism in Eastern Europe in the postwar period. Dissidents like Rudolf Bahro shaped an independent Marxist/Left politics that had an impact on cross-European debates through its engagement with environmental politics. Under state socialism, conservationists and nascent environmental activists worked under extremely difficult organizing conditions to bring the environmental problems of “actually existing socialism” into contestation. In Czechoslovakia, a movement for “conservation and environmental education” was formed in 1958, which led to some politicization of environmental issues. In Poland, “escape to nature” movements based on a critique of everyday life in the socialist city were prominent in cities such as Katowice, whereas other urban movements politicized the socio-ecological problems of the city. The mobilizing of
such environmental claims and politics played a central role in the struggles for democratization throughout Eastern Europe in the late 1980s and early 1990s. In the period of transition, however, the conditions for environmental nongovernmental organizations and campaigns have often been fraught with risks. In Slovakia, for example, movements such as Greenpeace Slovakia and Za Matku Zem (for Mother Earth) have faced harassment and criticism.

The negotiation of the diverse trajectories and political identities of the different constituents of the European green movement has often been the subject of contestation. Die Grünen has been by far the most successful green party in Europe in parliamentary terms and was involved in government in coalition with the German Social Democratic Party between 1998 and 2005. Despite this, the party’s parliamentary strategy was the subject of internal dissension between “fundamentalists” and “realists.” Parties such as Die Grünen have had to experiment with the organizational structures developed in extraparliamentary social movements and have had to
maintain support and links with diffuse constituencies. Green movements in other national contexts have faced similar challenges. Such practices of institutionalization can be contested and uneven. In the United Kingdom, recent debates over the Green Party’s decision to move from a shared to single leadership and concerns over a break with nonhierarchical structures and participatory democracy attest to the continuing character of such tensions. Despite such tensions, environmental movements have been adept at placing environmental issues and regulation on political agendas, including those of the European Union.

Bruno Latour has argued that a possible future for the green movement is for its concerns to become an uncontentious and consensual part of common sense. Left critics of the green movement have long argued that its early focus on issues such as population depoliticized environmental problems and ignored their social components. Such a depoliticization of political ecology continues to be significant. In the United Kingdom, a consensual construction of environmental issues was mobilized as part of the “Third Way” political agenda in the 1990s. What is significant, though, is that despite the institutionalization of many of their core concerns over the past 20 or 30 years, European green movements continue to generate contestation of unequal social and environmental relations in multiple ways. Green activists have been at the forefront of the struggles against neoliberal globalization. Through doing, so they have shaped the politics around “environmentalisms of the poor,” which link social justice and environmental issues, and they have outlined alternatives to neoliberalism, such as localized economic practices. Finally, radicalized and assertive nonviolent direct action continues to be an important part of green politics in Europe, most recently demonstrated by the climate change protests at Kingsnorth in the United Kingdom (the proposed site of a new coal-fired power station).

The European green movement was established out of diverse trajectories of political activity and has played out differently in different places. It has had a major impact in making environmental concerns central political issues of our time. The movement has brought such issues into contestation through multiple ways, and it will surely continue to be diverse and contested.

David Featherstone

See also Antisystemic Movements; Chipko Movement; Deep Ecology Movements; Environmental Social Movements; International Environmental Movements; Political Ecology

Further Readings


EUROPEAN UNION

The European Union (EU) was founded in 1992 in Maastricht, the Netherlands. It began functioning as a political organization in November 1993 and has gradually become one of the key actors in international affairs. In Europe, the EU has clearly become a political unit toward which the minor, small, and great powers tend to look before anything else and toward which regional, political, and economic activities increasingly tend to be geared. The EU currently includes 27 European states and is structured around three key institutions: the European Parliament (with 736 members), the European Commission (with 27 commissioners), and the European Council. It is through these political bodies that EU legislation—the driver of integration—is debated, passed, and implemented.
Evolution of the European Union

The EU is the most recent stage of the post–World War II European integration process, which was launched in the early 1950s. Over time, European integration has been legitimated by various dominating discourses. Throughout the process, it has been commonly argued that European integration is aimed at maintaining peace in Europe. From the 1950s up to the 1970s, integration was designed not only to restrain the spread of Soviet communism into Western Europe but also to hinder the rise of fascism.

The institutional dimension of this process has developed step by step during the past decades. The European Coal and Steel Community (ECSC), founded in 1951, first governed the coal and steel industries of the six founder states. In the 1960s, the customs union, a common agricultural policy, and the first seeds of the common market were established for the Community, which now included the ECSC, the European Economic Community, and Euratom. In the 1970s, the political institutions of the Community became stronger; since 1979, the members of the European Parliament have been directly elected every 5 yrs. (years). The late 1980s were a formative time for the development of the European single market, which finally took effect in 1992. After the end of the Cold War, European integration has been increasingly legitimized by the argument that without the EU, the member states would be too small to come to terms with the heightening international economic competition. The establishment of the EU gave more power to the Union, leading, for instance, to the subsequent monetary union based on a common currency (the euro), and a common foreign and security policy. The common currency became reality in 2002 with the adoption of the euro.

Even though the 27 member states all need to fully implement the acquis communautaire (the total body of EU legislation), the institutional bonds within the EU vary. For instance, all the member states are not included in the European Economic and Monetary Union (EMU). Only 16 members have entered the EMU and adopted the euro as their currency. Some older member states, such as Britain and Denmark, have stayed outside the “Eurozone,” whereas only three states of the 2004 enlargement (Slovakia, Slovenia, Cyprus, and Malta) have been included in the EMU. Moreover, Ireland and the United Kingdom remain outside the Schengen regulations, which aim to dissolve the physical borders within the EU, whereas some of the nonmembers, such as Norway, currently participate in this system.
The Nature of the Integration Process

There are several schools of thought in European integration studies that seek to explain and understand the successes and failures of the integration process. The early classical theories, such as federalism, were political models targeted toward containing the hostilities between France and Germany by limiting national sovereignty. This process involved rescaling the political power that regulated the strategically important economic sectors of these states. The classical models were followed by theories such as neofunctionalism and liberal intergovernmentalism, which usually emphasize not only political but also economic interdependence. The member states are willing to participate in the integration process because it is both politically and economically rewarding—at least for some segments of society. More recent theories, for example, the social constructivist approach, highlight the importance of both national identities and material interests in the successes and failures of the EU. The multilevel governance theory has also become a vibrant field that explains the complex rescaling of government and governance within the confines of the EU and beyond.

Despite the steady institutional development of the EU and its predecessors, it is often argued that the EU has remained an elitist project with little significance in the everyday lives of its inhabitants. There has been a lot of debate around the democratic deficit of the EU, indicating that the political decisions made by the EU take away too much decision-making power from member states, local governments, and their citizens. As a result, the EU launched a principle of subsidiarity to limit the powers of the EU and keep more political decision making at the local and national levels. One reason behind such a political move is that the emotional attachment of “EUropeans” to the territory of the EU has remained weak.

Seen from the perspective of political competencies, the EU is clearly more than an international
organization but less than a modern Westphalian state, whether federal or unitary. Even though it uses significant power over the member states in sectors such as agriculture and common trade policies, and even though its economy is now partly controlled by the European Central Bank, the EU does not have control over the armed forces and cannot make decisions about war and peace within its territory. The EU thus does not possess a monopoly on legitimate physical violence within its territory, as Max Weber might have put it. In other words, the EU uses significant political power within its borders, but its power has paradoxical limits. For instance, the recent declaration of independence by Kosovo reveals the somewhat amorphous nature of the EU. However, regardless of its significant contribution to the independence process, the EU does not have the ability to recognize Kosovo as a sovereign state.

One of the peculiarities of European integration is that both politicians and scholars tend to have a constant need to redefine the fundamental nature of the EU. At first sight, the EU does not seem to correspond to any previously known polities. As John Pinder stated in 1995, the nature of the EU as a political organization is difficult to understand as it escapes the conventional categories of international relations and federal state. It is sufficient to define the EU as a constant process that does not have self-evident political or geographical finalities. In essence, the institutional and territorial development of the EU is contested to the core and is therefore under serious political debate not only in the European Parliament, the European Commission, and the European Council but also within the member states. Often the territorial and institutional issues of the EU become part of political struggles within the member states, where political opposition is able to pressure government by using various rhetorical resources that integration provides.

In scholarly debates, the EU is frequently conceptualized as a new postnational polity based on multilevel governance. One of the most useful territorial approaches to multilevel governance is offered by Ole Jensen and Tim Richardson, who proposed in 2005 that the gradually evolving spatial policies of the EU form a European “monotopia”: a particular ideology that views the EU territory as one space that should be organized around large metropolitan areas and the logistic flows that connect these nodes. The emerging territoriality of the EU thus clearly has a particular spatial selectivity that is built into it.

The EU has also been conceptualized as a new type of imperial political subject. Ole Wæver’s treatment of the EU in 1997 as organized around concentric circles is a well-known attempt to look at Europe’s map through an alternative framework. He suggests that the EU is a particular type of empire that pulls other states into its orbit. Wæver points out that the EU is organized around one center and tries to keep this core intact. Simultaneously, the EU implicitly disciplines its immediate outsiders, the so-called close outsiders, using various techniques. Wæver also points out that beyond the circle surrounding the core, the EU also seeks to intervene directly in political development—without a significant military capacity.

The models that emphasize the EU’s imperial features should also take into account the often allegedly elitist nature of the European integration. It is widely accepted, for instance, that the EU expansion in 2004 was specifically based on attempts to Europeanize the political elites within the applicant states. One may also treat the politics of conditionally, the effective screening, monitoring, aiding, and gate-keeping practices of the EU institutions that seek to Europeanize the applicants, as a tool through which the EU specifically seeks to influence the various political elites within its neighborhoods. These observations come close to the ideas of John Agnew, who suggests that the EU is a form of hegemony as it has insinuated itself into everyday life not only within the EU but also in its neighborhoods and within its trading partners. Agnew thus defines the EU as a relatively light administration, a network of centers united around common goals and policies that can expand flexibly both in the scope of what it does and in the geographical area in which it operates without collapsing.

The Future of the European Union

All the previous enlargements have put significant pressure on the integration institutions and, therefore, reshaped the nature of the integration. In the history of European integration, territorial
Everglades Restoration

Over the past century, much of the Florida Everglades has been transformed from a principally subtropical wetland to a human-dominated system of agriculture and urban development. The present Everglades system is half its predisturbance size, the underlying water table is lower, the cyclical hydro-periods are altered, and the

Expansions have often been preceded by institutional reforms. The EU is currently preparing for future expansions. The negotiations specifically touch on the Balkan states and Turkey. Ash proposed some years ago that geography and the future of the EU were inseparable. He suggested that the kind of political community the political actors within the EU think the EU will be inevitably determines how far they want to enlarge the Union. And, conversely, how far they want to expand the Union determines what kind of political community the EU can be. The possible Turkish membership of the EU is particularly challenging in this context. There are those who reject the membership as they see it as jeopardizing not only the unity of the EU (or its mythologized and essentialized European values) but also the power structure within the EU. Some opponents script Turkey as a non-European state both geographically and culturally. However, Turkey’s EU membership enjoys considerable support within the EU—surprisingly also among federalists. Some of them perceive Turkey as an important strategic location that will stabilize the European geopolitical order.

The EU seems to be a crisis-prone organization that regularly faces serious challenges that jeopardize its very existence. In 2007, Jan Zielonka proposed that the EU suffers from a sort of midlife crisis and that coping with that crisis requires coming to terms with its increasingly diversified cultural, political, and economic space. He asserts that the postenlargement EU is not best described in terms of a Westphalian political unit and that the EU will not work as a Westphalian state. Zielonka thus suggests that the EU should develop in a direction that could be conceptualized as a neo-medieval empire—a sort of metagovernor that distributes decision-making competence between multitudes of territoriality and functionally defined self-governing actors. In such a view, the EU would not impose tight hierarchical control; rather, it would act as a mediator between various European networks and as a facilitator of continuous communication, cooperation, and compromise between such networks. With the neo-medieval empire, the borders of the Union should be flexible and open to those neighbors who embrace the basic set of liberal values and accept the rules operating within it.

European politicians often seem to demand more statelike development from the EU, especially when the process of integration encounters severe institutional difficulties or faces problems in affirming its political legitimacy among the people. Statism represents a particular mind-set according to which proper political life is possible only within statelike polities that are centrally organized, strictly bordered, and territorially homogenized and within which the people share a solid identity that attaches them to the bounded homeland. This form of territorial reasoning is still common within the EU, and the debates on the future of the EU are likely to be plagued by it.

Sami Moisio

See also Eurocentrism; Euromarket; European Green Movements; Globalization; North Atlantic Treaty Organization (NATO); Trade; Transnationalism

Further Readings


Over the past century, much of the Florida Everglades has been transformed from a principally subtropical wetland to a human-dominated system of agriculture and urban development. The present Everglades system is half its predisturbance size, the underlying water table is lower, the cyclical hydro-periods are altered, and the
vital freshwater sheet flow is diverted to the Atlantic Ocean and the Gulf of Mexico (see satellite images A and B). These anthropogenic changes have affected the Everglades by isolating and impounding the wetlands, reducing the wildlife habitat, degrading the water quality, and promoting the invasion of nonindigenous vegetation. In recognition of the imbalances created within the Everglades system by humans, the state of Florida is currently undertaking a multibillion-dollar restoration project as it attempts to repair the degradation. The Everglades Forever Act of 1994 mandated this restoration project to increase freshwater flow back into the Everglades and to establish minimum pollution levels for phosphorus runoff from the bordering Everglades Agricultural Area (EAA). This 40-yr. (year) restoration plan, approved by the U.S. Congress in 1999, is titled the Comprehensive Everglades Restoration Plan.

Historical Context

Large-scale modifications of the Everglades began with the Swamp Lands Act of 1850, when the state of Florida claimed ownership of more than 20 million acres of the Everglades with the plan to drain the system so as to attract private investment...
and economic activity. In 1880, the developer Hamilton Disston initiated the first significant drainage efforts designed to convert the wetlands into fertile agricultural land. In 1904, Napoleon Bonaparte Broward made the Everglades drainage the centerpiece of his successful, gubernatorial campaign as he invested significant political capital into drainage efforts that he promoted as creating vast, economic opportunities for Floridians. From 1906 to 1930, the Everglades experienced the construction of three major drainage canals—North New River, Hillsboro, and West Palm Beach—and completion of the Miami Canal.

To enhance drainage, a levee system was constructed in the 1920s around the southern shore of Lake Okeechobee, which prevented the natural sheet flow of fresh water from Lake Okeechobee south to Florida Bay. Concurrently, the Tamiami Trail highway was constructed to connect the rapidly developing communities on the eastern and western coasts of South Florida, which blocked nearly all surface water flow and divided the enormous, contiguous Everglades into two distinct northern and southern systems.

The second major phase in the human alterations of the Everglades began in 1948, when the Central and Southern Florida Project for Flood Control (C&SF Project) was authorized by Congress to expand greatly the previous drainage canal network for South Florida’s booming human population. The Central and Southern Florida Flood Control District, now known as the South Florida Water Management District (SFWMD), was given authority to manage the system, with the primary goal of preventing flooding in the rapidly growing urban areas. The major alterations of the C&SF Project included the construction of more levees and conversion of several hundred thousand acres of natural wetlands to the EAA.

The last large-scale Everglades alterations included a series of flood control levees added in the late 1960s and early 1970s to divide the remaining northern Everglades into three large water conservation areas (WCAs). The WCAs, although they are considered natural areas, are intensely managed systems that have little resemblance to their historical state. The primary purpose of WCAs is not environmental conservation but ensuring adequate water storage and supply for agriculture and coastal development.

Additionally, the water stored in the WCAs contains high phosphorus levels from EAA runoff, which is the main reason why invasive, nonnative plants have outcompeted the once dominant native vegetation in these systems.

### Comprehensive Everglades Restoration Plan

After a century of exploiting the natural system, the 1999 Comprehensive Everglades Restoration Plan (CERP) represents an ideological shift in environmental thinking for South Florida, acknowledging that Florida is facing the possible extinction of a unique and vital regional ecosystem. CERP’s fundamental hypothesis is that the reduction of environmental stressors such as drainage systems, flood control structures, and pollutant runoff will reinitialize the functional ecosystem processes, thus moving the system toward a more natural state. However, due to the widespread human impacts on ecosystems, it is unlikely that the system can be completely restored to its condition prior to human impacts.

To accomplish the Everglades restoration, CERP established three primary goals. First, and most important, is to restore the natural hydrology. The entire biotic structure of the Everglades system depends on reestablishing the natural water quality and quantity. Second is to restore the natural habitats, including historic populations of key flora and fauna (see photo). Third is to establish a sustainable compatibility between the natural ecosystems of the Everglades and the human-built systems surrounding it. To meet these goals, CERP proposes 68 projects. These projects can be grouped into seven general components that primarily address water storage, conservation, and delivery to both natural and urban areas:

1. **Surface water storage reservoirs**, both within the EAA and in the surrounding counties, will hold more than 1 million acre-feet of fresh water.

2. **Aquifer storage and recovery systems**, constructed around Lake Okeechobee, will store fresh water in deep wells for later recovery.
3. *Storm water treatment areas*, consisting of managed wetlands, will reduce pollution from both the EAA and urban runoff.

4. *Seepage management systems* will help raise the water table within the Everglades by reducing groundwater flow from the western natural areas to the eastern built areas.

5. *Freshwater sheet flow*, to the Everglades, will be reestablished by eliminating levees and canals.

6. *Rainfall-driven water management* will be used by the SFWMD to deliver water into the Everglades in a manner that replicates historical hydro-periods.

7. *Water reuse and conservation strategies* will address water quality issues for municipal wastewater discharge.

More than 8 yrs. after its adoption, none of the CERP projects has been completed. Significant programmatic advances have been made, however: four CERP projects are under construction, and several pilot projects are being evaluated. The primary reasons for CERP implementation not moving forward can be attributed to bureaucratic, regulatory, and budgetary impedances. Unfortunately, the delay in CERP implementation means that the Everglades system continues to degrade as the human population and its impacts continue unabated in South Florida.

_Tobin Hindle_

*See also* Conservation; Environmental Management; Environmental Restoration; Environment and Development; Landscape Restoration; Population, Environment, and Development
As a basic definition, the term everyday life refers to those ordinary, taken-for-granted, habitual thoughts, activities, and settings that are close and familiar to all of us but that are rarely measured by governments or scholars or endowed any particular significance. Henri Lefebvre used the metaphor that everyday life is like fertilizer: It functions as a source of life-giving power, but it largely goes unnoticed as it is tramped underfoot. According to Lefebvre (1991), “a landscape without flowers or magnificent woods may be depressing for the passer-by but flowers and trees should not make us forget the earth beneath, which has a secret life and a richness of its own” (p. 87). Lefebvre also famously coined the phrase “the familiar is not necessarily known” to capture the understanding that while an activity such as shopping or walking may be ubiquitous, it lacks meaning if it is not recognized—by naming, counting, researching, and assigning value to it.

Not only does the essence and substance of the life we live each day largely go by without us consciously thinking about it, but when we do turn to scrutinize ordinary, individual moments, taking these fragments out of their background context, they tend to assume an unrealistically distinct, abstract, and spectacular status. This highlights a tension that has recently emerged in academic scholarship, between descriptions of everyday life as myriad phenomena to be recorded and interpreted from diaries, biographies, ethnographies, and the like; theories of the everyday relating to personal, emotional, and normative compulsions; and the politics of everyday choices and behavior with respect to production, consumption, and social reproduction. For instance, the decision to buy or boycott a particular brand of good or to shop with a reusable shopping bag assumes a symbolic political function in affluent societies today.

Accepting that everyday life is a contested concept, John Eyles implores us never to treat it as unproblematic or unimportant. This plea to name and critically consider the sum of forces and events that create and sustain human existence owes much to efforts by feminist economists to demonstrate that everyday life has to be reproduced—socially and materially. Feminist scholars have long recognized the moral and political significance that what is most frequently omitted from, or trivialized within, the formal economy are the tasks of unpaid, home-based work, largely performed by women, representing all that it takes to keep individual workers, families, and human societies fed, clothed, housed, educated, and cared for. For example, in a book called The Secret Life of Cities (alluding to Lefebvre’s metaphor above), Jarvis and colleagues conjoin the concept of everyday life with that of social reproduction to examine them both as a category of practices and as unconscious knowledge supporting the reproduction of human society—and the biographic techniques of analysis required to bring this issue to public attention.

Recent years have witnessed the idea of everyday life entering the popular imagination. For instance, the phrase “part of everyday life” is used to advertise public transport in the English city of Newcastle. Across the social sciences, the term has been variously adopted to convey a shift
in scale (to the body, household, street, and locale), greater emphasis on qualitative methods, and a reconnection of humanistic geography with the temporal rhythms of age, history, calendar, and clock. As with the coinage of any umbrella concept (e.g., sustainability or globalization), there is potential for precision of meaning to be lost to the illusion of an all-encompassing understanding of the everyday as everything, all the time, everywhere. The definitions provided here identify parameters key to the use of this term.

**Origins of the Concept and Language of Everyday Life**

By defining everyday life with reference to the reader, along with the author, the concept and language of everyday life is located squarely within a humanistic geographic tradition. Analytically, the term is variously used with reference to the urban sociology of Georg Simmel, the rhythm analysis of Henri Lefebvre, the dramaturgy of Erving Goffman, the phenomenology of Michel de Certeau, the theory of practice of Pierre Bourdieu, the feminist institutional ethnography of Dorothy E. Smith, and the time-geography of Törsten Hägerstrand.

One body of work through which the concept and language of everyday life can be traced far back in time is that of symbolic interaction and an “emotional” or “therapeutic turn,” stimulated in the 1920s by Freudian psychoanalysis. Interest in the people’s compulsion to come together through face-to-face encounters, along with the gendered performance of “emotion work” that this entails, is reflected in the seminal work of Erving Goffman’s dramaturgical approach, elucidated in the 1959 publication *The Presentation of Self in Everyday Life*. Parallels are drawn between everyday life and the staging of a drama, with reference made to a “front-stage” and a “back-stage” “performance” of social activity. The enduring significance of personal, social interaction is associated with growing awareness that people reflexively monitor their own and other people’s emotions in a dynamic process of “learning to fit in.” This reflexivity and the skills required to “pass” as “normal” in a given situation assume particular significance in an age where it is both possible and increasingly expected that “virtual” communications will replace “real” social contact. Feminist scholars have recently extended this analysis to consider other neglected, less visible, highly gendered “infrastructures of constraint,” including, for example, guilt, love, and obligation. Some commentators argue that these “ties that bind” (and the self-interest of the family as an institution of social organization) disadvantage women as mothers, who are in effect creating a public good (well cared for, socially adjusted youngsters) for intrinsic reward in the absence of pay.

A second body of work traces a shift in emphasis from the extraordinary and the exotic to the ordinary and the mundane. It has been used by cultural geographers to legitimize the study of marginalized groups and discriminatory practices on grounds of gender, class, age, ethnicity, and family form. Caution must be applied in the renewal of interest in “ordinariness.” Just as “everyday life” can appear to encompass everything, everywhere, so too the notion of ordinariness can appear slippery or culturally value laden. Renewed interest in “ordinariness” and “everyday life,” as illustrated in the work of Walter Benjamin and Georg Simmel, for instance, interprets ordinariness through the voyeuristic eyes of the male flâneur, the gentleman stroller of the city. Similarly, the 1960s’ situationist international movement drew attention to the “spectacle” of the everyday, where “situations” were consciously staged as spectacular events intended to disrupt the “alienation and suffering” of capitalist social interaction. The aim of the situationists was to create freedom for the individual through transgressive roles and situations.

When feminist scholars evaluate the writings of Simmel, Benjamin, de Certeau, and the Situationist International on subjects of gender and the family, they reveal the construction of a masculine understanding of the world in which the everyday realities of women and minority groups and the exploitation they frequently experience in the home or street do not register in the philosophical imagination. In contrast, notions of ordinariness have been mobilized by feminist urban planners whose objective is to make the material fabric and institutions of the city as inclusive as possible. Clara Greed, for instance,
draws attention to the inadequate provision of public toilets. She highlights the way poor design and lack of investment in the built environment restricts the everyday mobility and access of older people, mothers with young children, and people with disabilities.

**The Geographies of Everyday Life**

The academic disciplines of geography and history are crucial to understanding everyday life as a world that we experience directly, through locally proximate, coordinated interaction. The spatial and temporal rhythms of everyday routines and practices were first represented in humanistic time-geography as a theory and method of tracing daily routines as coordinates of time and space. Seminal work by Torsten Hägerstrand portrayed mundane “projects” of daily life, such as shopping for groceries on the way home from work, as a function of multiple “pockets of local order,” the path or trajectory of which is governed by three contextual constraints relating to capability, coupling, and authority. A striking image associated with this work is a three-dimensional prism, mapping an individual’s path across a morphological map and relating this to a 24-hour daily time budget. Allen Pred described this sequencing and timing as the choreography of existence. More recently, Anthony Giddens developed the theory of structuration to capture and resolve the mutually co-constitutive structures and agency of individual and collective time-space relations.

Apart from time and space, a humanistic geography tradition highlights the importance of an integrated analysis of people-place relations. This
is demonstrated in the work of EuroFEM, a loose affiliation of European feminist scholars committed to promoting and implementing gender mainstreaming in planning policy, job creation, and local initiatives. The EuroFEM model for implementing change is one of participatory engagement, smoothing the barriers to everyday coordination by drawing on historic ideals of collective living.

Renewed interest in everyday life has coincided with innovative methods and techniques to communicate the rhythm of daily life; its seasonal and economic cycles; the dynamics of age, aging, and the life course; and the importance of local history, customs, and language. Methods include the collection of daily diaries and biographies/autobiographies and ethnographic and participatory research. There is a long tradition of looking at time budget analysis in relation to household responsibilities, the provision of child care, and more recently the commoditization of these activities by transnational paid domestic workers. Thus, while everyday life emphasizes micropractices, it also captures the local imprint of global flows and circuits of people and cultural and economic practices.

Today, new advances are being made in the application of geographic information systems and geovisualization to conventional methods of in-depth ethnography. Sophisticated satellite tracking devices can now be applied to the complex problems of everyday life to log not only the time-space coordinates of activities and social interactions but also perceptual experiences felt along the way (fear, discomfort, intimidation, landscape aesthetic, and the like). These innovations suggest a more integrated and nuanced analysis of the social reproduction of everyday life in the future.

Finally, although the term everyday life has taken on particular significance in the 21st century, as a result of the “cultural turn” in geography, it is important to note that interest in mundane observations is not new. Back in 1937, mass observation was established in Britain specifically to study the everyday lives of ordinary people. It continues today to focus on different and diverse aspects and views of daily life with a newly commissioned e-mail communications archive. Other projects in the archive include the National Panel of Diarists, the Bolton Worktown Project, Pub Conversations, and the Cooperative Correspondence Club of young mothers.

Helen Jarvis

See also Behavioral Geography; Chicago School; Class, Geography and; Commuting; Cultural Turn; Discourse and Geography; Emotions, Geography and; Ethnicity; Existentialism and Geography; Feminist Geographies; Gender and Geography; Hägerstrand, Torsten; Home; Humanistic Geography; Identity, Geography and; Positionality; Pred, Allan; Situated Knowledge; Spaces of Representation/Representational Spaces; Structuration Theory; Time-Geography

Further Readings


The term existentialism was coined by Jean-Paul Sartre, the most famous of the existentialists, to refer to a range of mid-20th-century European philosophers who studied the conditions of humanity: Sartre, Albert Camus, and Gabriel Marcel in France; Martin Heidegger, Karl Jaspers, and Martin Buber in Germany; and Ortega y Gasset in Spain, as well as myriad novelists and playwrights who promoted a historical genre based on the “existential condition.” These thinkers shared a belief that established philosophical frameworks were insufficient to understand or improve human conditions. The basic premise is that human beings bring their existence into the world or become human by engaging a world peopled with other humans. A common referent is Sartre’s idea that “existence precedes essence,” that is, we make ourselves through our human projects.

There are two significant strands of existentialism in geography: (1) existential phenomenology, associated with Heidegger, and (2) existential Marxism, associated with Lefebvre and Sartre. Existential phenomenology arose during the 1970s to contest the spatial determinism of positivist science by emphasizing the subjectivity of human becoming. Profoundly influenced by the Heideggerian concept of *dasein* ("being in the world"), the project of recentering the human subject was addressed by scholars such as Ann Buttimer, Yi-Fu Tuan, and Edward Relph, who advocated a “humanistic” geography that linked phenomenological openness to the project of being with the social and cultural acts of creating, giving meaning to, and sustaining places or landscapes. Buttimer emphasized human creativity and intersubjectivity in everyday, experiential geographies, or “life-worlds”; she held human values as a basis for establishing how people care about one another. Yi-fu Tuan sought to understand how being open to the creative possibilities of human beings affects interactions with the surrounding environment and the struggle to make their landscapes. Edward Relph used phenomenology as a basis for assessing the ways in which humans design and give meaning to valued landscapes, emphasizing the connection between care for landscape and care for humanity.

Other scholars have made a direct link between a phenomenology of being and a spatial ontology. Marwyn Samuels drew on Heidegger and Sartre to construct a “biography of landscape” as the physical expression of human spatial relations, looking to Martin Buber to understand spatial relations as the project of overcoming the distance between people and to suggest that the recursive act of setting others at a distance and entering into relations is informed by shared belief systems, or worldviews, that define the qualities and normative values of human societies. The inevitable distance that separates humans is thus actively filled with the project of becoming-for-others. For John Pickles, such a spatial ontology defines the scientific paradigm through which we create geographical knowledge.

The ongoing intellectual influences of these humanistic geographers are profound, including the incorporation of meaning and values firmly within the discipline; highlighting questions of ethics in geography, both as a subject of research in understanding human actions and as a normative framework for conducting research; and laying the methodological groundwork for the discursive turn during the 1990s, which views landscapes as repositories of human meaning and spatial relations as centered on the individual and as constituted through human action through which (to paraphrase Derek Gregory) landscapes are made to mean.

Existential Marxism came later to human geography. During the 1970s and 1980s, humanism and (structural) Marxism were viewed as oppositional. Humanist geographers objected to what they saw as economic determinism, and Marxist geographers viewed existential phenomenology as romanticist, individualistic, apolitical, and unable to address the complex problems of political economy or the role of power in defining human relations. Since the poststructuralist turn, however, with its emphasis on the active, social construction of meaning, existential Marxism, with its roots in Hegelian rather than Heideggerian dialectics, has had a profound effect. Most notable is the
influence of Henri Lefebvre, who has inspired geographers not only with his dialectical materialism but also with his concept of “space” as a social production. Existentialist in his concern to overcome the rupture in traditional philosophy between theory and practice, Lefebvre posited a triad of representations of space, representational space, and spatial practices (perceived, conceived, and lived space), which has had wide application by geographers in understanding urban processes and, in particular, the question of how “public space” is created and occupied. Don Mitchell, for example, draws on Lefebvre’s notion of the “right to the city” as a starting point to explore the relationship between spatial production, human rights, and class-based locational conflict.

Although he has received negligible attention from geographers, Sartre’s existential Marxism holds great potential for overcoming a tendency to read Lefebvre through an instrumentalist lens—and thus to reify “space,” as well as to extend a spatial ontology. Notwithstanding the many divisions among the French intellectuals that persisted over time, Lefebvre and Sartre shared the designation of existential Marxists, but their interests took different trajectories by the 1970s, when Lefebvre turned his attention to the relationship between cities and rural areas (hence his influence on urban geographies), whereas Sartre was concerned with the social relations of colonialism and oppression, influencing, among others, the work of Franz Fanon and Edward Said. Sartre’s project was to go beyond the individual to understand history as an “accumulation outflow” of spatialized acts and the spatiality of human existence as a dialectically constitutive process of “spatializing-spatialized,” through which human beings become objects to/for one another in a vast social geography of systematically constructed conditions. Using Marxian concepts such as scarcity, seriality, and alienation, he argued that colonialism and capitalism are the major structuring systems that, paradoxically, both situate the possibility of humanity and create alienation, thus preventing humans from achieving their full humanity. Sartre’s quest to overcome history as human oppression is relevant today for understanding how ongoing spatialized systems of human relationship contribute to the effects of colonialism, Othering, and racialization.

Audrey Kobayashi

See also Buttimer, Anne; Critical Human Geography; Ethics, Geography and; Humanistic Geography; Phenomenology; Relph, Edward; Sense of Place; Tuan, Yi-Fu

Further Readings

Exotic species, also termed nonindigenous, non-native, alien, naturalized, or introduced, are species that are not native to a particular geographic area and were introduced intentionally or unintentionally by human activities. These introductions allow the species to overcome natural barriers to dispersal. In practice, classifying a species as native or exotic can be difficult when it is unclear how or when a species was introduced or if hybridization occurred with native species. Exotic species can be any type of organism, including microorganisms, plants, or animals. It is important to recognize that a species exotic in one area has an area to which it is native and that the behavior of the species in the two habitats can be completely different. Key features of exotic species include their geographic distribution, threats to introduction areas, and amenability to control treatments.

The pervasiveness of human movement and activity means that exotic species are present throughout the world. Even the Antarctic Peninsula is home to an exotic spider crab (Hyas araneus) and unknown numbers of exotic microorganisms. Exotics typically become distributed by long-distance dispersal (e.g., moving from one continent to another by ships or planes), followed by local secondary dispersal (e.g., transportation by cars or animals from one area to another). Multiple introductions can be required for a species to become established, and a slow initial population growth may be followed by exponential growth as invasion proceeds. A classic example of an exotic species, the European rabbit (Oryctolagus cuniculus), invaded Australia in this way. Repeated introductions made over a 70-year period beginning in the late 1700s failed, but when establishment did occur in 1859, populations grew exponentially to devastate agricultural crops expansively.

Exotic species vary in their impacts on introduction areas. Some exotics become inert components of their new habitats or provide replacement functions for native species that have become extinct. Many exotics benefit humans and are often purposely introduced (e.g., medicinal plants, food species). Overall, however, exotic species are considered the second largest cause (behind direct habitat loss) of native species extinctions. Exotics also disturb human activities and cause billions of dollars of damage annually. Exotic mussels, for example, clog municipal water intake pipes, reduce the utility of beaches, and alter fisheries.

Competitiveness, invasiveness, and the ability to engineer ecosystems are characteristics that make exotic species threats to native ecosystems. Exotics often have competitive advantages in their new environment because they escape the agents that had kept their populations in check. For instance, the prevalence of the Asian bark fungus (Cryphonectria parasitica) remains low in its native habitat because host trees have attained resistance. When this fungus was inadvertently introduced to North America on imported trees near the turn of the 20th century, it eliminated the nonresistant American chestnut (Castanea dentata) as an overstory tree.

Many exotics are invasive, readily colonizing and dominating habitats, especially disturbed areas containing open niches available for colonization. Exotics also often invade and thrive in high-resource environments (e.g., nitrogen-enriched soils due to pollution) more than native species do. The ability to invade and usurp...
Resources quickly provides exotics with further competitive advantages over native species.

Species that can “engineer” ecosystems by drastically changing disturbance regimes or habitat conditions for indigenous biota are some of the most devastating exotics. Annual grasses of the genus *Bromus*, for instance, have invaded the arid lands of Western North America, engineering a new disturbance regime (frequent fire) by providing fuel to ecosystems where sparse fuel historically limited the spread of fire. Native communities, not adapted to frequent fire, are destroyed, while the exotic grasses exploit this new disturbance regime.

Land and water managers can attempt several techniques to limit the establishment of undesired exotics, starting with preventing dispersal (e.g., by cleaning ships). If prevention fails, surveys and monitoring can detect new occurrences of exotics. Control treatments are most cost-effective and biologically feasible when exotics are in the early phases of invasion and have low abundance.

Suitable control treatments for exotic species depend on factors such as population sizes and traits of the exotic species, restrictions on the use of some treatments (e.g., herbicides), and economic costs. There are two general classes of treatments: species focused and habitat focused. Species-focused treatments act directly on individuals and include poisons, shooting, physical removal (e.g., hand pulling of plants), genetic engineering (e.g., introducing sterile individuals), and biocontrol (introducing another organism, often derived from the exotic species’ native habitat, to control the target species). Biocontrol is risky, and there are examples of the biocontrol organisms themselves becoming troublesome exotic species. Habitat-focused treatments involve manipulating the habitat to be less favorable to the invader. Examples of these techniques include filling potential colonization sites by seeding native plant species, altering soil characteristics to render it outside the limit of an invader’s tolerance, or implementing a disturbance regime (e.g., grazing) unfavorable to the exotic. A key condition of habitat-focused treatments is that the treatment should harm the exotic more than resident native species.

There are some dramatic examples of complete eradication of an exotic species, especially when the species threatens agricultural crops. Generally, however, exotics are continuing to expand their populations. Improving our understanding of which exotic species will have severe impacts to enable prioritization of treatments, how to reduce dispersal of exotics and invasibility of potential habitats, and species-specific control treatments is important for effective exotic species management.

*Scott R. Abella*

See also Biodiversity; Biogeography; Environmental Restoration; Human-Induced Invasion of Species; Invasion and Succession; Land Degradation; Pest Management
**Further Readings**


---

**EXPLORATION**

The history of geographic knowledge is closely bound up with explorations. Exploration is a means by which the geographical horizons and imaginations of cultures and peoples are extended, bringing the unknown world into the known.

Historically, explorations required curiosity and initiative as motivating forces, as well as endurance and force to stand up to the perils that harsh climates, hostile landscapes, and rough waters present; but above all, they require a courageus temperament to face the unknown. As centuries went by, new aims were added to the list of the reasons for exploring other regions and continents: economic and commercial goals, political conquest, or scientific missions that allowed people to acquire a deeper familiarity with the planet. Nowadays, virtually no place on Earth remains unexplored.

When explorations involved territorial conquests, they typically involved territories that were already inhabited by other populations. Explorations under these circumstances often turned into brutal aggression toward the locals. For example, the principle of *Terra Nullius*, a Latin expression meaning “nobody’s land,” was applied during the 16th century by European conquerors in many colonized areas of the world and neglected the rights of indigenous peoples. Despite those tragedies, however, adventurous travel and mapping of new territories greatly contributed to the development of geographic knowledge.

The history and geography of exploration make up a complex tale in which power and economic and military imperatives are interwoven with the biographies and (often colorful) personalities of individuals. Although the term seems archaic today, *exploration* has long been central to the formation of geographic knowledge. This entry offers a general overview on geographic explorations, focusing on some of the most fascinating protagonists and events.

---

**Prehistoric Exploration**

The history of geographic explorations is as ancient as humanity itself and is often entwined with migrations that crossed vast parts of the Earth. Early humans moved to new territories in order to hunt animals or find a more favorable climate and landscape. About 70,000 yrs. (years) ago, human ancestors in Central Africa often faced terrible droughts. To find new sources of food and sustenance, they moved toward the eastern coasts of the continent, learning how to succeed in surviving in different contexts. They essentially followed two directions: One group, thanks to lower sea levels due to the ice age of that period, succeeded in reaching Southern Asia and, millennia later, went ashore in Australia, settling islands in the Pacific Ocean, or continued toward Central and Northern Asia until they found a natural bridge that allowed them to set foot on the American continent; others covered a shorter distance and arrived at Europe after crossing Mesopotamia. Simply by walking, a mode of travel that allowed them to cross a few hundred miles per year, humans exited Africa around 75,000 BC, spreading to East Asia and Europe by 50,000 BC and crossing the Bering Straits into North America between 20,000 BC and 30,000 BC. Boats and the technology for conquering maritime space were particularly well developed among the early Polynesians, who dominated the Pacific Ocean by AD 300, when they reached Easter Island. The final chapter of the diffusion of people to the planet’s remotest regions was the Lapita complex of cultures of the
Bismarck Archipelago, which developed the double-hulled outrigger canoe, an innovation that allowed travel out of sight of land and thus initiated the waves of Polynesian expansion that ended with the settlement of Hawaii around AD 1300.

**Ancient History**

Around the Mediterranean Sea, several cultures showed an interest in the discovery of new sites. After exploring their main source of sustenance, the Nile River, ancient Egyptians focused on exploring the eastern part of the African continent, in particular the Horn of Africa. By 2500 BC they reached what they considered the legendary Pwenet or Land of Punt, where precious materials such as metals and timber, as well as slaves, could be found.

According to tradition, in the 10th century BC, the Phoenicians, who roamed across the Mediterranean, traveled from the Red Sea to Southwestern India and Sri Lanka. Four centuries later, they may have circumnavigated Africa by passing the so-called Pillars of Hercules, which were widely considered to be the end of the world. They also reached the archipelago of Madeira, the Azores, and the Canary Islands, and once in the Atlantic Ocean, they reached Cornwall in the north.

The Greek exploration of the Mediterranean Sea and Asia contributed greatly to the creation of maps and to the rise of cartography. However, they focused their attention on exploring other regions as well. Around 350 BC, for instance, the geographer Pytheas arrived in Britain and Norway. He described, for the first time, the frozen sea and some of the natural phenomena found in the northern parts of the world, such as aurora borealis and the midnight sun, and also provided social and cultural observations. He also described Thule, a mysterious land that scholars later identified either as Iceland or as the Norwegian coast but that at the time of Pytheas, and later throughout the Middle Ages, symbolically represented the northernmost border of the world.

Alexander the Great was a prominent figure in the Greek panorama of military explorations. Geographers and topographers followed him during his travels, and that greatly contributed to the knowledge diffusion of the new conquered world.

In the map later drawn by Eratosthenes of Cyrene, in fact, we can notice that the center of the world was no longer Greece but Persia, which was the core of the empire built by the Macedonian.

The enormous empire created by the Romans between the 1st century BC and the 4th century AD was due to their insatiable thirst for power, which led them to build a dominion that embraced virtually all the Mediterranean countries. Roman space was enlarged by repeated wars of imperial conquest, including the Punic Wars, which annihilated Carthage, and subsequent campaigns that added Gaul, Britain, Thrace, Greece, Syria, Palestine, and Egypt. The term *orbis terrarum* reflected the widespread view that the borders of the empire coincided with the habitable world, the boundaries between order and chaos. However, the Romans enjoyed robust trade with India and even occasional interchanges with China.

**The Vikings**

During the 8th and 9th centuries, the Vikings played a great role as explorers. On board a type of longship called *drakkar*, they penetrated Russia and the Mediterranean and reached Normandy, Britain, Ireland, Iceland, and a land they called Greenland. Renowned for attacks throughout the coasts of Europe, they perfected the keel by adding the steerboard (from which comes *starboard*), a side rudder that stabilized larger ships when hauling commodities. Their brief arrival in America nearly 500 yrs. before Columbus is a much discussed issue: Various arguments support different, and often contrasting, hypotheses about the European discovery of the continent. Some ancient texts and the results of archaeological excavations in the northwest coast of Newfoundland and the Labrador Peninsula led some scholars to support the theory of Viking journeys from Greenland to the region that the Norse explorer Leif Eiriksson named “Vinland,” “the land of grapes” or “the land of wine,” that is, Newfoundland. Around 1950, a map was discovered, presumably drawn in 1440, which gave new support to the theory of the Norse discovery of America (Figure 1). Besides Europe, Africa, and Asia, it represents the northern islands of Greenland and Iceland and a territory named Vinlanda Insula, which was said to have been discovered in the
Figure 1  The Vinland map, a 15th-century world map purportedly based on a 15th-century original
11th century. However, some recent studies have indicated that the map is false as it was drawn using recent ink. The debate about the authenticity of the Vinland Map continues.

Marco Polo

Early explorations linking Europe and Asia typically centered on the Silk Road (a name coined by the geographer Ferdinand von Richthofen in 1877), the umbilical cord that connected Europe, the Middle East, Central Asia, South Asia, Tibet, and China with ceaselessly flowing caravans of goods, innovations, ideas, merchants, missionaries, and armies. Few individuals traveled the entire distance of the Silk Road; rather, it was served by networks of intermediaries. For 2,000 yrs., its camel caravans formed the primary artery of commerce, linking ports, trading cities, oases, and innumerable different cultures. The Venetian Marco Polo, precapitalist history’s most famous adventurer, as noted in his account *The Discovery of the World*, was the eminent figure of Western exploration in the 13th and 14th centuries.

Marco Polo was one of the first Europeans to arrive in China, which he called Cathay. He belonged to a rich family of merchants and began his famous voyage with his father and uncle in 1271. They went to Persia with the aim of embarking to India but then continued their journey until they reached Kashgar in Western China. They went across Hindu Kush and were the first Europeans to reach Pamir. Then they skirted the Taklimakan Desert and arrived in Lop Nuur. After crossing the Gobi Desert and 3 yrs. after their departure, they reached the court of the Mongol emperor Kublai Khan in Khanbalik, now Beijing (Figure 2).

Arab Explorations

The rapid expansion of Islam from the 7th to the 12th centuries unified diverse cultures under the umbrella of a common religion and, inasmuch as Arabic came to dominate governance and religious life in North Africa and the Middle East, a common language. The Arab world’s situation intermediate between Asia, Africa, and Europe gave it a strategic location within the enormous trade networks that linked together places as far-flung as China, Mozambique, and Belgium. Arabic dhows plied the Persian Gulf and the Indian Ocean for centuries, and Arab shipbuilders invented the mizzen mast, the lateen sail, and stern rudder, all of which shortened times at sea. Unlike the Red Sea, which was notoriously difficult to navigate, the Persian/Arab Gulf emerged as a navigable waterway crossed by innumerable dhows. By the 9th century, when the mechanics of the monsoons (from Arabic *mausim*, “season”) had been deciphered, Arab trade became regularized across the Indian Ocean.
The most famous Arab explorer, Ibn Battuta, born in Tangier in 1304, was a Berber explorer who traveled throughout the Islamic world and wrote a chronicle of his journeys, a precious source of the geographical knowledge of the medieval era. He started his voyages with the ritual pilgrimage, or hajj, to Mecca and went on to visit new places for the next 30 yrs., covering more that 70,000 mi. (miles) from Spain to Western Africa and from the Middle East to India and China. During his journey back from Timbuktu, he met nomadic Tuareg people and lived several years with them, reporting on their culture and social structure.

Chinese Discoveries

By far, the most famous example of Chinese exploration occurred during the Ming dynasty. It was, of course, the famous expeditions led by Cheng Ho, one of many royal eunuchs, who made seven voyages across the Indian Ocean between 1405 and 1433 as an emissary for the Ming emperor. The “treasure ships” of the Chinese navy, at 7,800 tons, were three times as large as anything the British had in the water prior to the 18th century. With 317 such vessels and a crew of 37,000, Cheng Ho made Malacca the headquarters for visits to southeast Asia and then went on to Bengal, Persia, Aden, and Zanzibar. In 1407, he conquered Annam. In 1414, he brought a giraffe to the Ming emperor, which caused a flurry of excitement, a sign that the African kingdom of Malindi had become a tributary state. In scouring the Indian Ocean for pirates, the treasure ships provided security and lowered the costs of transport. However, China’s expansionism was short-lived. Japanese pirates subjected Chinese ships to unending harassment. Eventually, Ming China suffered multiple social economic crises that, coupled with the intense internecine battles between conservative Confucian bureaucrats and ambitious eunuchs, ended Cheng Ho’s voyages when all coastal shipping was banned in a fit of xenophobic nationalism. Such moves also reflect the Ming dynasty’s redirection of state resources from foreign expansion to internal security. The Ming withdrawal from the Indian Ocean by the 1430s created a power vacuum later filled by the Portuguese.

The European Age of Discovery

From the beginning of the 15th century till the late 19th century, as capitalism spilled across the world’s oceans and continents, Europeans were very active in the field of exploration. The invention of the sternpost rudder in the 12th century allowed for easier travel in the stormy waters of the North Sea and the Baltic; travel here had been difficult for earlier Greek or Roman ships, which relied on steering oars. Late medieval navigation relied on tables of declination, books of sailing instructions, the astrolabe and quadrant, and a vast oral system of instructions. The astrolabe, long known to Arab and Indian sailors, was introduced into Europe in the 11th century and was designed to solve trigonometric calculations in long-distance navigation. Similarly, feudal Europe was greatly affected by the introduction of the magnetic compass, the first mention of which in Europe dates to 1190; its subsequent adoption facilitated and catalyzed the process of early Western expansion, reducing the uncertainty of travel over both land and water and allowing sailors to sail the Mediterranean during winter. The greatest technical achievement of medieval navigation was the ability to sail against the wind, which was accomplished using sails on the new carack vessels, which were more sophisticated than the square sails used by the Romans. The result was an increase in the velocity and magnitude of trade across the Mediterranean.

In 1453, the Ottomans conquered Constantinople, which caused serious difficulties for non-Muslim merchants who dealt in goods from Asia. The search for new routes to China and India became enormously important for the powerful merchant class in Europe; moreover, innovations in the process of sailing and in cartography allowed great explorers to obtain extraordinary results.

In 1488, the Portuguese captain Bartolomeu Dias rounded the Cape of Good Hope, and 9 yrs. later, the navigator Vasco de Gama reached the Indian Ocean, sailing until he reached India. In 1497, the Italian John Cabot explored Newfoundland. And in the early 16th century, Amerigo Vespucci proved that what Columbus believed to be Asia was in fact a continent previously unknown to the Europeans. Ferdinand Magellan
EXPLORATION

The Arctic and the Antarctic

The North Pole and the South Pole, with their extremely rigid climate and hostile waters, always represented a formidable challenge for explorers. In 1909, the American explorer Robert Edwin Peary claimed to have been the first person to reach the North Pole, although in 1989 the National Geographic Society stated that he stopped at 5 mi. south of it. Sir Walter William Herbert is now generally considered to be the first person to have achieved that feat in 1969, after having spent years exploring the region.

At the southernmost end of the world, the South Pole was the focus of a competition between the Norwegian Roald Amundsen and the British Robert Scott, who died during his return journey after having reached the South Pole a couple of weeks later than his adversary.

The Northwest and the Northeast Passages

At the beginning of the 20th century, exploration continued for the long sought after Northwest Passage, the sea route that allowed one to reach the Pacific from the North Sea through the Bering Strait. The fact that the Indies could be reached by sailing west from Europe was known shortly after Columbus’s famous journey. New routes to reach Asia soon became the priority of merchants and explorers. In 1497 and 1498, John Cabot was the first to seek the Northwest Passage, but the attempt was extremely difficult because of the general lack of knowledge about the area and the perils hidden by ice.

At the beginning of the 17th century, two English navigators accepted the challenge: In 1611, Henry Hudson reached the strait and the bay that were named after him, and William Baffin, in the same period, conducted five voyages in the Arctic Ocean. He became the first European to discover the bay that now bears his name. The passage was navigated for the first time by the Swedish explorer Adolf Erik Nordenskiöld in 1879. Due to the difficult environmental and climatic conditions involved, the dream of reaching the passage was abandoned for some time, but in the 19th century, the Royal Geographic Society charged the Scottish voyagers John Ross and William Edward Parry with

succeeded in navigating the straits near Tierra del Fuego, and his famous voyage across the Pacific in 1520 remarkably encountered no storms, thus giving the ocean its name. Although he died in the Philippines, his ships ultimately completed the first circumnavigation of the globe. Magellan’s crew’s feat was not only technically significant, but the very act of global circumnavigation gave rise for the first time to the notion that there was a closed, finite world, an early inkling of the relative spaces yet to come.

Likewise, the European discovery of the Americas initiated a momentous material and discursive transformation. Columbus’s narratives of the discovery were drenched in medieval cosmological meanings, including claims that the New World represented Paradise, or heaven on Earth, the site in which saved Christians would enter as promised by Genesis, or, equivalently, a newfound Jerusalem. Spanish imaginary geographies, for example, portrayed the Atlantic as an extension of the Mediterranean and the conquest of the New World—the decisive spatial fix of the 16th century—as an extension of the reconquista long after the Moors had been expelled from Iberia. The ideological impacts of the discovery of the New World, as its implications reverberated across Europe, are difficult to exaggerate. The so-called shock of the primitive not only fueled notions of racial hierarchy and European superiority, they also prompted heated debates as to what constituted a human being. O’Gorman (1961) suggests that the discovery of the New World accelerated the incipient secularization of Western society; in his account, the Americas were not simply discovered but discursively invented, a process that obliterated the medieval Christian worldview of a planet with three, and only three, continents and accentuated the division between humans and nature (Figure 3).

As the Enlightenment unfolded, European explorations also served as a means of generating vast quantities of data, in the form of samples of plants and animals as well as scientific observations. The most famous example of scientific exploration was perhaps Captain Cook’s voyages to the Pacific in the mid 18th century, which included botanists, naturalists, and scientists studying the transit of Venus and searching for the mysterious terra australis.
Figure 3
Historical map of the Old World, as Ptolemy knew it, by Johannes de Anshein, Ulm, 1482
Source: http://commons.wikimedia.org/wiki/File:World_of_Ptolemy_as_shown_by_Johannes_de_Anshein_—_Ulm_1482.png
exploring the northern coastline of America. After discovering the north magnetic pole, Parry explored the Arctic by sled, learning from the Inuit about how to survive in such extreme conditions. But it was again Amundsen who succeeded in crossing the passage during the voyage he undertook from 1903 to 1906.

**Explorations in Africa**

Africa, of course, was always known to Africans. Its discovery by others, however, is a different matter. The ancient Greeks, Phoenicians, and Romans made numerous exploratory voyages to the African coasts and established numerous colonial settlements there. When in the middle of the 15th century the power of Islam became an obstacle for European trade with Asia, merchants tried to find alternative ways to get to the Indies, including circumnavigating the African continent; however, its interior territories still remained mostly unknown and mysterious, imagined to be inhabited by fantastic and monstrous creatures.

In 1788, the African Association for Promoting the Discovery of the Interior Parts of Africa was founded, which later became the Royal Geographic Society. Its main aim was the exploration of the Niger River and pinpointing the location of the ancient city of Timbuktu. Among the European explorers of African regions was Simon Lucas, who toured Southern Libya. In the 19th and 20th centuries, European colonists were attracted to Africa by its abundant resources and commercial potential. The most famous explorer of the Dark Continent was David Livingstone, a Scottish missionary who dedicated his life to travels in its southern area, starting in 1845. He succeeded in crossing the Kalahari Desert and in reaching Ngami Lake, in Northwestern Botswana. Between 1852 and 1855, he explored the Zambesi River and found the Mosi-oa-Tunya falls, which he named Victoria, in honor of the Queen. He later discovered Lake Niassa (also known as Lake Malawi) and the Tanganyika Lake. Livingstone, like many explorers, focused his attention also on the quest for the Nile’s source, a question that had aroused the curiosity of geographers since ancient times. The solution of that enigma was finally found by the British adventurers Richard Francis Burton (1821–1890) and John Hanning Speke (1827–1864).

**Women Explorers**

Although they remain too often in the shadows of a male perspective on history and geography, many women can be found in the panorama of explorers. Due to their social marginalization, women explorers have always had to fight against prejudices, sometimes even hiding their identity by disguising themselves.

During the first half of the 19th century, the Austrian Ida Pfeiffer traveled all around the world. Accustomed to independence and physical activity since she was a child, she departed, by herself, and eventually met the indigenous tribes living in the rain forests in South America, whom she considered to be primitive and inferior to Christians. Among her most interesting adventures are the journeys through Mesopotamia and the jungles of Borneo: Such destinations were considered by some to be impossible for a woman traveling alone. Despite her election to the geographical societies of Berlin and Paris, she was not admitted by the British Royal Geographical Society because she was a woman (Figure 4).

Born in Geneva at the end of the same century, Isabelle Eberhardt traveled through the Algerian desert, dressed as a man and pretending that her name was Mahmoud Essadi. She was charmed by the huge open space of the desert and attracted by the Arab culture and way of living.

At the same time, Mary Henrietta Kingsley was interested in Western and Central Africa. She left Britain and was the first European to explore isolated areas of the French Congo and Gabon. During her travels, she did not stop despite numerous difficulties, learning from the locals how to survive in the jungles, discovering new species of fish, and not even hesitating to study cannibal tribes.

In Asia, the Belgian-French writer and explorer Louise Eugénie Alexandrine Marie David, whose pseudonym was Alexandra David-Néel, satisfied her spiritual thirst by traveling first through India, where she studied Buddhism and even lived for
several years in a cave, and reaching Tibet in 1924, thus being the first European woman to enter Lhasa.

Susana Servello

See also Blaut, James; Bowman, Isaiah; Cartography, History of; Colonialism; Columbus, Christopher; Cook, Captain James; Enlightenment; Eratosthenes; Eurocentrism; Gama, Vasco da; Geographical Imagination; Historical Geography; Human Geography, History of; Humboldt, Alexander von; Ibn Battuta; Ibn Khaldun; Imperialism; Jefferson, Thomas; Lewis and Clark Expedition; Magellan, Ferdinand; National Geographic Society; Orientalism; Pilgrimage; Poles, North and South; Portolan Charts; Postcolonialism; Powell, John Wesley; Ptolemy; Race and Empire; Royal Geographical Society; Russian Geographical Society; Travel Writing, Geography and

Further Readings

EXPLORATORY SPATIAL DATA ANALYSIS

Exploratory spatial data analysis (ESDA) is an extension of exploratory data analysis (EDA) to detect and understand the properties of spatial data. ESDA is an important research field in geographic information science (GIScience) because we now live in a data-rich world and a vast and increasing amount of information on geographical phenomena is recorded digitally as spatial data. Adopted by many subdisciplines of geography and other disciplines (e.g., epidemiology), ESDA has made enduring contributions to the development of GIScience and geography. Therefore, ESDA is an important topic in this Encyclopedia. This entry discusses three related topics: (1) EDA, (2) traditional ESDA, and (3) the potential future directions of ESDA.

Exploratory Data Analysis

Tukey, who introduced the notion of EDA, thought of it as the first step in a two-step process similar to criminal investigations. The first step focuses on evidence collection, and the second step—the confirmatory analysis—evaluates the strength of the evidence. The difference between the two steps is similar to that between descriptive and inferential statistics: The former produce data descriptions, and the latter draw inferences by assuming some statistical models. Employing descriptive methods, EDA imposes few prior assumptions about the data and attempts to let the data speak for themselves. EDA focuses on understanding trends and outliers rather than on fitting models. Good defines EDA as a collection of descriptive techniques that summarize data properties, detect patterns and unusual features in data, and formulate hypotheses from data. EDA techniques integrate descriptive statistical methods (e.g., median, mean) with visualization methods (e.g., charts, graphs, and figures). Modern EDA emphasizes human-computer interaction. It employs highly interactive, dynamic statistical graphics that allow the analyst to manipulate the views of the data directly in a simple, intuitive way. The views are often dynamically linked, so that the selection on a subset of data in one view can be immediately reflected in all the other views. Such data manipulation allows investigation of data from various perspectives concurrently and exposing insights that are otherwise hidden.

Exploratory Spatial Data Analysis

Spatial data are special. According to Tobler’s First Law of Geography, “everything is related to everything else, but near things are more related than distant things.” In spatial data, an observation is referenced to a spatial entity (e.g., a point, line, or areal unit). Because the phenomena represented by neighboring entities interact with each other, samples of spatial data are usually not independent. The dependence in spatial data, often referred to as spatial autocorrelation, conflicts with the basic assumption of independent observations in statistics, making traditional statistics and most EDA methods invalid for spatial data. For example, EDA methods estimating regression parameters and generating measures of fit become invalid in the presence of spatial autocorrelation. Extending the notion of EDA, ESDA methods are geared especially for exploring spatial data—summarizing the spatial properties of data, detecting spatial patterns in data, and generating hypotheses based on the patterns. ESDA techniques consist of a collection of spatial statistics and traditional visual methods including geographic maps.

The special characteristics of spatial data determine the goals and methods of ESDA. For EDA, Tukey (1977, p. 208) modeled any data value with two components:

Data = Smooth + Rough,
EXPLORATORY SPATIAL DATA ANALYSIS

where “Smooth” refers to the “predictable” part (i.e., fit) and “Rough” is the “unpredictable” part (i.e., residuals). Haining (2003) extended the Tukey model in spatial terms:

\[
\text{Spatial data} = \text{Spatial smooth} + \text{Spatial rough},
\]

where “Spatial smooth” includes spatial trend (i.e., the systematic variation of an observation along a region, such as a mountain) and spatial autocorrelation (i.e., in the case of positive autocorrelation, neighboring areas are more similar in attribute than distant areas). This component results from and reflects spatial dependency. The spatial rough component includes localized hot or cold spots and spatial outliers (i.e., places whose attributes are very different in magnitude from the values in their neighboring locations); this component reflects spatial heterogeneity (i.e., the spatial difference and nonstationarity associated with each location). Traditional ESDA techniques focus primarily on detecting and visualizing (a) spatial distribution, (b) spatial association (e.g., autocorrelation and overall clustering), and (c) local clusters and hot spots.

First, visualization of spatial distribution is concerned with displaying the values of a variable observed in a region and exposing the spatial trends and outliers. A simple way to identify spatial clusters is to link a boxplot to a map. The value distribution of an attribute is displayed along the boxplot; data items outside the upper and lower “fences” of the boxplot are considered as outliers and can be highlighted on the map. Smoothing methods are often employed to reveal general trends in spatial data, especially for maps consisting of many small areas. A simple numerical form of smoothing is the spatial average, which averages the original value of a place and the values at its surrounding places. The key issue of map smoothing is to improve the precision of the statistic while not introducing much bias. Spatial averaging based on more surrounding places can increase precision (i.e., with increased sample size), at the cost of introducing more bias since a place “borrows” values from areas farther away.

Second, the focus of ESDA on spatial association primarily involves two aspects: (1) global autocorrelation testing overall clustering (i.e., the null hypothesis states that data values are distributed randomly across geographical space) and (2) local autocorrelation detecting local clusters. Generally, spatial autocorrelation can be measured from two perspectives, lattice and geostatistical, based on the way in which spatial proximity is conceptualized. Anselin (1994) interpreted the two perspectives in a more intuitive manner: neighborhood view and distance view.

The neighborhood view, more commonly found in geography, conceptualizes a place comprising discrete points or areal units and spatial neighbors as spatial units that have a common boundary. The neighborhood is formalized in a spatial weights matrix \( W \), with elements \( w_{ij} = 0 \) when locations \( i \) and \( j \) are not neighbors and \( w_{ij} \neq 0 \) otherwise. A location has spatial interaction only with a countable number of its neighbors (when \( w_{ij} \neq 0 \)). The global autocorrelation can be measured by a number of methods including Moran’s \( I \), Geary’s \( C \), and \( G \) statistics. The basic idea is to compare the observation at each location with the average of its neighbors. If most locations have the observation similar to the average of observations at their neighbors (e.g., a high-value location is surrounded by neighbors with high values), a positive spatial association is suggested, otherwise a negative spatial association is suggested (e.g., a high-value location is surrounded by neighbors with low values). Global spatial autocorrelation can be visualized using the Moran scatterplot (Figure 1). A scatterplot plots a variable value against its spatial lag for each location. A spatial lag of a location is defined as the weighted average of observations at the neighboring locations; the weighted average is calculated using the weights in the spatial weights matrix. The method follows the interpretation of Moran’s \( I \) statistics as a regression of a location’s spatial lag on the variable itself. The regression line through the points on the scatterplot reflects the Moran’s \( I \), thus indicating the spatial autocorrelation (e.g., a 45° slope indicates strong positive spatial autocorrelation). In addition, local spatial autocorrelation can be detected through the outliers in the scatterplot. The outliers do not follow the same process of spatial autocorrelation of the other points and can be considered as locally nonstationary. A more sophisticated tool is the LISA map (Anselin, 1995). GeoDa, a freely downloadable software application, implements the methods mentioned above.
The *distance view* is based on geostatistics. The approach considers spatial observations as a sample of points from a continuous spatial distribution. It conceptualizes spatial interaction as a continuous spatial process. Therefore, spatial autocorrelation is formalized as a continuous function of distance. The global autocorrelation can be measured by comparing the observation dissimilarity (e.g., square difference) at a pair of locations with the spatial distance between the two locations. The autocorrelation can be visualized by a variogram, in which the differences between all pairs of locations are sorted by spatial distance.

Finally, detecting local clusters is important in many contexts, such as epidemiology and crime studies. A local cluster is an unusual concentration of events in an area, where counts or rates are considerably larger than expected values based on an assumption of spatial randomness. Although descriptive statistics methods (e.g., Local Indicators of Spatial Association [LISA]) help identify local hot spots and outliers, many researchers employ inferential statistical methods for detecting clusters. One popular method is the scan statistical method. One of the earliest versions of the spatial scan statistical method was the Geographical Analysis Machine (GAM) of Openshaw and colleagues. Built on the concept of GAM, Kulldorff’s spatial scan statistic gained popularity in recent years because it is both deterministic (i.e., it identifies the locations of clusters) and inferential (i.e., it allows testing of the statistical significance of the clusters). The method assumes the null hypothesis that events are randomly distributed in geographic space. The alternative hypothesis is that there are an unusually high (low) number of events within a subarea as compared with the outside areas. Kulldorff’s spatial scan statistic is implemented in publicly downloadable software—SaTScan. Because SaTScan does not provide cartographic output, users need to export SaTScan output into geographic information system (GIS) software to display the geographic location and size of the identified clusters (Figure 2). Chen, MacEachren, and Lengerich (2008) recently developed a visualization technique to facilitate exploratory interaction with SaTScan clusters and tune the clustering parameters.

**Potential Future Directions**

Computational methods can also be combined with visualization methods to support ESDA.
Computational methods that are of relevance to ESDA include machine-learning and data-mining methods supporting automatic identification of useful structures, patterns, and clusters in data. Examples of such methods include classification trees, cluster analysis, and neural networks. In recent years, research efforts in ESDA have been extended beyond univariate spatial data to include temporal, multivariate, and moving data as well. Large data volumes and the high multidimensionality of data sets, containing tens of variables, make the traditional ESDA methods no longer satisfactory. Hence, there is a need to develop integrated and more robust visualization and computational methods. Examples and detailed discussions on this topic can be found in Ferreira de Oliveira and Levkowitz (2003); Chen, MacEachren, and Guo (2008); and Keim, Panse, Sips, and North (2004). Another developing research field that is relevant to ESDA is geovisual analytics (GA). Conceived as an extension of visual analytics (Thomas & Cook, 2005), GA goes beyond ESDA to give more attention to support for both computational and human analytical reasoning processes. GA aims to facilitate a human analyst in his or her direct interaction with visually represented information in order to gain insights, draw conclusions, and ultimately make better decisions (see Keim, Mansmann, Schneidewind, & Ziegler, 2006). GA integrates techniques of visualization, human factors, statistics, mathematics, and data mining.

Jin Chen

**Figure 2** Visualization of high-risk clusters of cervical cancer in the United States (2000–2004) identified by SaTScan

Source: Map created by author, based on data from the National Cancer Institute using the Surveillance, Epidemiology, and End Results (SEER) program via the SEER*Stat software.

Note: Orange = high risk, white = normal, blue = low risk.
EXPORT-LED DEVELOPMENT

Export-led development is a phrase encompassing a number of different economic development strategies—especially export-oriented industrialization (EOI) and the export of primary products—pursued by the developing world. These practices arise from development theories informed by economic and political economic thought. A number of major and conflicting lines of economic and political economic thought, including neoclassical economics, development economics, and Marxist political economy, hold different views on export-led development. Disagreement centers primarily on the role of the state in guiding economic growth through export orientation and the desirability of pursuing comparative advantage versus competitive advantage. Geographers have been prominent in theorizing the uneven development of capitalism and the successes and failures of various nations’ attempts at economic development.

Richard Peet makes an important distinction between “strong” development, in which the productive resources of society are used to help the poorest people, and “weak” development, characterized by economic growth that provides more for everyone but primarily benefits the elite and is supposed to “trickle down” to the poor. Beyond this distinction, the concept of development has been subjected to a number of important critiques from the perspectives of sustainable development and poststructuralism. Those who agree that development should be pursued measure it by various criteria, including gross domestic product/gross national product, the Physical Quality of Life Index, and the Human Development Index.

Models and Theories

There are several perspectives on export-led development, which offer differing views on its causes and consequences.

Further Readings


Structuralism

The structuralist perspective, informed by both development economics and critical political economy, including dependency theory and world-systems theory, has viewed state intervention as both useful and necessary to industrialize developing nations and thereby reduce global inequalities. The heterogeneous theories critical of laissez-faire and neoclassical economics generally agree that a lack of state intervention in the economy means that developing countries will not move beyond primary production and simple manufacturing without state intervention and, therefore, that international trade exacerbates international inequalities. In particular, structuralists have pointed to declining “terms of trade”—the value of a country’s exports relative to its imports—for primary product exporters. The declining price of primary products relative to manufactured goods means that primary product–exporting countries must export ever-higher volumes to purchase the same quantity of manufactured goods over time. Additionally, underdevelopment may be entrenched by the “resource curse,” which points to the strong but paradoxical relationship between high stocks of resources such as minerals and oil and low levels of economic development.

Import Substitution Industrialization

After World War II, a structuralist consensus formed in developing economics and became the dominant theory guiding economic policy in developing countries. In the 1950s, most developing countries pursued import substitution industrialization (ISI) to encourage domestic manufacturing to meet domestic needs. ISI drew on the experience of Latin American nations that had pursued these policies since the Great Depression as a solution to export dependency. Under ISI, governments imposed tariffs on imported manufactured goods to limit competition (but did not impose tariffs on inputs for domestic manufacturing) and provided incentives and assistance to the local firms that manufactured these goods. Fiscal policy involved raising the exchange rate of domestic currency, thereby making imports without tariffs cheaper and making exports more expensive (which favored investment in domestic industry).

Nations such as Brazil, the paradigmatic case for ISI, promoted domestic industries through support for domestic firms, state ownership of some sectors, and allowing entry to foreign firms and transnational corporations that the state required to be oriented toward meeting domestic needs. Brazil by the 1970s met the majority of domestic demand for steel, textiles, transportation equipment, and pharmaceuticals. Large nations, with potentially large domestic demand, generally fared best under ISI, while the domestic demand of smaller countries was too small for efficient mass production. But there were several problems with the ISI model. First, ISI did not reduce inequality within the countries that pursued it. Second, ISI did little to reduce imports, as the sectors targeted by ISI required considerable inputs of capital goods such as machinery. This strategy in many instances led to large national debts, which, together with the interest rates hikes of the late 1970s, ushered in the debt crisis of the 1980s.

Export-Oriented Industrialization

The other development model often contrasted with ISI is EOI, although it is important to note that the countries most strongly associated with EOI also pursued ISI, often at the same time. EOI sought industrial growth and structural change in national economies—away from primary products and simple manufactured goods toward higher-technology sectors—by nurturing the conditions domestically that would allow exports to compete successfully in international markets, rather than relying only on domestic demand as promoted under ISI. The newly industrialized countries (NICs)—especially Taiwan, South Korea, Hong Kong, and Singapore—are considered the paradigmatic cases of EOI. They shifted from a largely ISI-oriented approach toward an EOI-oriented approach in the late 1950s and early 1960s. The South Korean and Taiwanese cases, together with their predecessor in EOI, Japan, and their successor, China, have shown that directed state intervention, including what neoclassical economists would call blatantly anticompetitive policies, have created some of the most rapid growth rates and structural changes to national economies that have ever been seen. These countries’ economic growth was strongly
supported by their efforts at exporting manufactured goods, but strong government intervention in prioritizing economic sectors is certainly not in line with neoclassical economics’ preferred laissez-faire approach but instead draws on the prescriptions of state intervention promoted by structuralism and other state interventionist thought.

The Heckscher-Ohlin Model

Within the realm of neoclassical economics, and in direct contrast to structuralist and radical political economic thought, the Heckscher-Ohlin model is hegemonic and forms the theoretical basis for global “free trade.” Advocates of this model argue that countries should integrate into the world economy by focusing on their comparative advantage—that is, they should specialize in the production and export of goods that they have in relative abundance while importing goods that are intensive in factors of production in which the country is relatively poor. Like laissez-faire arguments at the end of the 19th century, neoclassical economics holds that state intervention in markets is harmful. Neoliberalism draws on neoclassical theory to oppose interventionist economic theories and has been hegemonic in its shaping of economic development in both industrialized and developing countries over the past 30 years. Export-led growth promoted by neoclassical economists and neoliberals eschews the state-led development trajectories that are common to EOI, such as prioritizing sectors by directly or indirectly directing investment, or state-owned industries. Thus, for most developing countries, the neoclassical and neoliberal prescription for national economic development is to continue to focus on primary products, in countries with relative wealth in natural resources, and labor-intensive, simple manufactured goods such as textiles because of the low wages prevalent in these countries.

Outcomes of Development

Those promoting neoliberalism have claimed that the experiences of the NICs support the prescription of pursuing comparative advantage. Thus, since the 1980s, the World Bank and the International Monetary Fund (IMF) have drawn on the successes of the NICs as the basis for structural adjustment as advocated by neoliberalism. With the debt crisis of the 1980s, the World Bank and the IMF imposed structural adjustment policies in return for continued access to loans. Structural adjustment generally required governments to reduce tariffs and quotas on imported goods, privatize state-owned sectors, cut government spending on public services to balance budgets, reorient their economies toward export sectors based on comparative advantage, prevent the overvaluation of exchange rates, create positive interest rates in real terms, and stop the subsidization of capital-intensive industrialization. Structural adjustment has proven to be highly controversial, with particularly fierce opposition among farmers in the developing countries, including members of the Via Campesina movement, who argue that they cannot compete with subsidized agricultural products from the United States and Europe and, as such, are being forced into new forms of exploitation reminiscent of colonialism. The threat of being cut off from the world economic system and the loans that allowed their governments to function proved too much for many developing countries during the debt crisis. Although most developing country governments were often hesitant and resistant, they had to implement structural adjustment programs as designed by the World Bank and the IMF.

In contrast to EOI, structural adjustment has primarily favored two strategies for export-led growth. First, export processing zones closely match the prescriptions of neoclassical economics in that they exploit developing countries’ comparative advantage in low wages (labor costs) by focusing on labor-intensive but low-skilled production—such as the textile industry and assembly of simple electronics—rather than attempting to pursue competitive advantage by changing comparative advantage through state intervention. Second, structural adjustment has prioritized exports of primary products as these are generally the resources in which developing countries have a comparative advantage.

The production and export of primary products—including goods from agriculture, forestry, fisheries,
and mining—have therefore flourished under the hegemony of neoliberalism. Export agriculture, often mirroring the large-scale plantations established under European colonialism, has boomed in developing countries to feed industrialized nations’ demand for year-round fresh produce. Export agriculture is typically critiqued as being more along the lines of industrialized agriculture in the industrialized world, meaning that it is intensive in its use of pesticides and fertilizers and results in numerous negative environmental and social consequences. Indeed, a large literature exists on the ecological and social crises arising from agroexport-led development.

Chile is used as the paradigmatic case of increased agricultural and timber exports under neoliberalism, beginning with Augusto Pinochet’s coup against Salvador Allende. Pinochet and the “Chicago boys,” trained in the neoliberal theories of Milton Friedman and others at the University of Chicago, made Chile the world’s first neoliberal experiment in 1973. The economic changes involved the liberation of prices and markets, encouragement of foreign investment, the destruction of labor’s power, minimization of environmental regulations, and a general commitment to “free trade,” making Chile extremely favorable for foreign investment. The goal behind the radical change in policies was to have capitalist enterprises and exports become the driving force for growth. Chile’s fruit production for export increased 40-fold, from US$30 million in 1974 to US$1.2 billion in 1996. As a result of this very rapid growth, Chile became the leading exporter of fruit in the Southern Hemisphere, servicing markets in both developed and developing countries. Critiques of the Chilean fruit boom include its reliance on the synthetic pesticides and fertilizers of industrial agriculture and the fact that the benefits are primarily distributed to the capitalist classes, whereas fruit workers have received little of the productivity gains and most can only find seasonal employment in the industry.

Ryan E. Galt

See also Colonialism; Comparative Advantage; Competitive Advantage; Debt and Debt Crisis; Dependency Theory; Developing World; Development Theory; Economic Base Analysis; Export Processing Zones; Foreign Direct Investment; Gross Domestic Product/Gross National Product; Import Substitution Industrialization; International Monetary Fund; Political Economy; Neoliberalism; Newly Industrializing Countries; Regional Economic Development; Structural Adjustment; Sustainable Development; Trade; Transnational Corporation; Underdevelopment; Uneven Development; World Bank; World-Systems Theory

Further Readings


An export processing zone (EPZ) is generally set up by a government to promote and capture foreign direct investment, international trade, technological transfer, and industrial development. These zones provide preferential treatments and incentives to attract foreign enterprises and investors. EPZs are also known as free trade zones (FTZs), special economic zones (SEZs), free zones (FZs), industrial estates, free ports, urban enterprise zones, and other terms.
The concept of EPZs originated in Spain in 1929. In the 1970s, EPZs became a global phenomenon as a developmental tool in low-income nations in Latin America, the Caribbean, Asia, and, to a lesser degree, Africa. By the 1980s, almost all developing countries had established at least one such zone.

The normal provisions are infrastructure, communication and financial services, elimination of trade barriers such as tariffs and quotas, exemptions from production and trade-related taxes and business regulations, friendly and simplified bureaucratic requirements and procedures, elimination of labor laws, and relaxation of environmental protections. Tax holidays, duty-free exports and imports, and unfettered repatriation of profits are regularly offered to foreign firms. EPZs are locations for foreign enterprises and nodes of international transportation.

The main objective for the host nations of EPZs is economic gain. These zones are the instruments to increase foreign currency earnings, upgrade production technology, enter the world market, and provide employment. With foreign investment concentrating in these zones, a multiplier effect to locally linked industries and a dispersal effect to the surrounding regions are expected. Innovation in international transportation and communication in the 1960s and 1970s allowed transnational corporations to lower production costs by outsourcing parts or all of the production to offshore locations. The division of the industrial production process, with corporate headquarters at home nations and production branches overseas, is the basis for EPZs. These zones are designed to absorb international branch operations. Globalization gives rise to selective international economic integration, and EPZs are the transnational connection points in developing nations. The flourishing of EPZ development since the 1970s is a spatial response to globalization.

Critiques

Developing states embrace EPZs as a developmental tool. Even though there are sufficient successful examples to attest to its effectiveness, there are, nevertheless, many counterarguments. In general, the multiplier effect and regional effects of many EPZs have been weak. The cost-effectiveness of government direct investment and indirect costs in relationship to the revenue received has been poor. Negative externalities have been reported, especially environmental pollution.

One of the most frequently raised issues of EPZs as an instrument of national development strategy is the competition among EPZs themselves. Since transnational firms seek economically advantageous sites for outsourcing and branch locations, and there are many suitable and available EPZs in the world, firms tend to seek
and bargain for the most advantageous zone. EPZs are forced to give their best offers to outbid their rivals. This competitive process is not only limited to international rivalry but also generates intensive contest between EPZs within the same country.

The most controversial issue is labor practice. International firms relocating the labor-intensive production segments are often opposed to labor unions and minimal-wage regulation. To be competitive, EPZs tend to avoid enforcing these institutional practices. Working conditions are substandard, and wages are low. Most of the employment is for young women in “dead-end jobs.” Often, EPZs do not generate new jobs but replicate the existing local jobs. Foreign firms, with better resources than the local firms, compete successfully at the expense of local industries for the “best” workers.

**Developmental Trend**

Since the 1990s, the World Trade Organization (WTO) has required member nations to eliminate import quotas and trade barriers, which EPZs offer. Export subsidies, which government supplies to the EPZ firms, are also to be repudiated. These WTO conditions have two effects on the member states: (1) making the entire national space more conducive and effective for export-oriented direct investment and (2) making EPZ establishment more difficult and economically less effective.

During the same period, many transnational corporations made a strategic shift from direct investment in low-value-added and labor-intensive segments of production abroad to sub-contracting, licensing, and franchising overseas firms. These new transnational tendencies move significantly away from the division of the industrial production process, the basis of EPZ establishment.

With both trends converging at the end of the 20th century, the demand for EPZs has been significantly reduced, and the development of EPZs has subsided.

Reginald Yin-Wang Kwok

See also Comparative Advantage; Competitive Advantage; Economic Base Analysis; Export-Led Development; Foreign Direct Investment; Globalization; Incubator Zones; Import Substitution Industrialization; Industrialization; Neoliberalism; Newly Industrializing Countries; Regional Economic Development; Trade; Transnational Corporation; Underdevelopment; Uneven Development

**Further Readings**


**EXTERNALITIES**

Broadly speaking, externalities (also known as “spillovers” or “neighborhood effects”) refer to “uncompensated welfare impacts,” that is, the actions and events of one party or person that affect the welfare (positively or negatively) of another without some type of remuneration. These arise when a decision maker does not reap
all the rewards or bear all the costs of his or her actions and can occur in both the production and the consumption of goods and services.

Positive externalities improve the welfare of an individual or group without a cost. For example, if one’s neighbor has an attractive garden or plays music that one enjoys, the receiving party derives benefits without paying the costs. Most positive externalities are relatively trivial. However, network externalities, which reflect the rising utility of systems such as telephone networks or the Internet, are important: The more users use a system, the greater the value it has to each user.

Negative externalities, however, which diminish the welfare of a person or a group, are a different story. Examples of negative externalities include the reduction in real estate values created by the location nearby of an unwanted land use (e.g., a toxic waste plant). If a developer erects a high rise that blocks a home owner’s view, the latter suffers a negative externality. More general cases involve the creation of air and water pollution, acid rain, noise pollution, or traffic congestion. Because the producers of negative externalities do not have an incentive to worry about the impacts of their actions on others, they generate social and market inefficiencies.

Negative externalities occur when the social costs of an action are not captured in the private costs in the form of the market price and are thus a prime example of market failure. For example, the true costs of operating an automobile include its impacts on highways, the environment, and public health, few or none of which are included in the price of gasoline or car insurance. A logging company may deprive a neighborhood of shade. In this case, and similar ones, the social costs are greater than the sum of the individual costs and lead to the overproduction of goods with high social costs. These are often seen as violations of individual rights and lead to serious ethical and political problems. Negative externalities are thus commonly cited as necessary instances of government intervention, such as zoning ordinances, health and safety regulations, and environmental conservation.

Geographers study the spatial location, frequency, and magnitude of negative externalities, which are unevenly distributed. The presence of a sports stadium, for example, may generate a field of noise that affects local residents negatively. These factors are particularly important in the analysis of transportation (e.g., congestion), land use (e.g., rural-to-urban land conversion), and natural resource conservation (e.g., hydroelectric dams and their impacts).

Barney Warf

See also Agglomeration Economies; Development Theory; Economies of Scale; Economies of Scope; Economic Geography; Knowledge, Geography of; Knowledge Spillovers; Location Theory; Neoliberalism; Not in My Backyard (NIMBY); Regional Economic Development

Further Readings

EXTINCTIONS

Earth is now entering its sixth period of mass extinction, defined as a major decrease in the number of species in a relatively short period of time. Extinctions are a normal part of the evolution of ecosystems. New species evolve, other species die out, and over millennia the diversity of life on the planet has increased.

The previous five mass extinction events are believed to have been caused by prehistoric meteor impacts and/or significant climatic changes (themselves caused by a variety of factors). The fifth event occurred 65 million years ago and wiped out an estimated 50% of Earth’s species, including the dinosaurs. In the current time, a similar percentage of Earth’s species are again threatened with extinction but this time as a result of the actions of the human species. The main causes include (1) habitat loss, (2) overexploitation, (3) introduced species, and (4) extinction cascades. Through these combined factors, around 70 species are believed to die out every single day; the rate of extinction is approximately 100 to
1,000 times greater than normal levels. In the coming years, anthropogenic climate change is also expected to add increasing pressure to struggling species and ecosystems, further accelerating the rate of extinctions.

Although extinction is a global event, the geographical distribution of extinctions is patchy; critical zones with high rates of extinction are unevenly distributed around the world. In addition, as the 2005 Millennium Ecosystem Assessment points out, the human experience of extinction effects falls disproportionately on rural communities and the poor. A geographical perspective brings a regional specificity to this global event and highlights the interconnections that link and amplify the extinction process. Four brief case studies will draw out connections between the diverse causes of extinctions and provide examples of the way in which a spatial analysis deepens our understanding of the loss and disturbance experienced when species die.

**Human/Nonhuman Entanglements in Place**

Even before a species becomes extinct, the role that it plays in an ecosystem may be severely impaired by dwindling numbers (called *functional extinction*). A good example is the case of Gyps vultures in India and the surrounding regions. Over the past decade, approximately 95% of these vultures have died as a result of poisoning caused by eating cattle treated with an anti-inflammatory drug called diclofenac. With vultures in decline, the millions of cattle carcasses they traditionally “cleaned up” have become a new source of food for growing populations of stray dogs and others. The increase in dog numbers is believed to be a significant cause for the spread of rabies and other diseases to humans, domestic animals, and potentially to some other threatened species. In addition, the vulture deaths are affecting people whose cultural and economic lives are directly connected with the birds, for example, the Parsi community, which traditionally exposes its dead to be consumed by vultures in “sky burials,” and “bone collectors,” who gather vulture-cleaned cattle bones for the fertilizer industry.

**The Ecological and Cultural Place of the Species**

The loss of “keystone” species has the potential to reduce the long-term viability of whole ecosystems and to reduce the capacity of other species to adapt to changing conditions, potentially leading to “co-extinctions.” For example, the spectacled flying fox (*Pteropus conspicillatus*) lives in the tropical world heritage rain forests of Queensland in Australia and is the major pollinator for the rain forests. These animals are currently under threat from habitat clearance and

An Oriental white-backed vulture (*Gyps bengalensis*) that was brought injured to the Sanjay Gandhi National Park stands inside a cage at the park in Borivili, a suburb of Mumbai, India, September 23, 2008. Several vulture species in Asia have been driven to the brink of extinction after eating cattle carcasses tainted with an anti-inflammatory painkiller given to sick cows. The drug has led to global population declines as high as 99% in slender-billed and other vulture species, especially in India.

*Source: AP Photo/Gautam Singh.*
the resulting disturbances, including invasive weeds. The loss of flying foxes endangers an ecosystem that is itself of great conservation value. With climate change, the role of these pollinators in maintaining the gene flow within plant species will be increasingly important. Flying foxes also remind us that the human experience of extinctions is local and enmeshed with cultural understandings. For some Australian Aboriginal people, flying foxes, along with numerous other extinct and endangered species, are totemic kin and are the focus of sacred songs, sites, and rituals. As one elder put it, “They’re not just fauna, they’re family.” In this context, local species losses entail the pain of losing nonhuman kindred whose existence in the world gave human life much of its meaning.

**Enclosed Spaces, Global Connections**

In the midst of so much loss, many of the Earth’s threatened species will survive only in zoos or other conservation centers. Survival is (perhaps) guaranteed inside these spaces, but many of the movements and relationships that organisms were once involved in will no longer be possible; their own lives and those of numerous others may be impoverished. For example, the Svalbard Global Seed Vault, in Norway, draws seed from around the world to provide a “backup” for many of the world’s agricultural plant varieties (which, alongside a great deal of “wild” biodiversity, are disappearing at an alarming rate). But seed in a bank (like animals in a zoo) is not equivalent to real beings living “in situ.” From inside these enclosed spaces, organisms are not able to be a part of the ecosystems they once inhabited, adapt to changing conditions, or be part of human communities. While these issues do not completely undermine the value of these conservation projects, they do highlight the need for critical thought about the enclosed and sometimes privileged spaces within which a great deal of conservation takes place.

**Interrupted Spaces**

Because the movement and genetic diversity of plants and animals are so important to the health of both individuals and the species, barriers that interrupt these movements are of vital importance. For example, the European bison (*Bison bonasus*) survives in forests in Poland and Belarus, but the herds are divided by a huge fortified wall. In this “interrupted” environment, fears are raised about the viability of the gene pool and thus about the future of European bison in the wild. In the context of conservation work, barriers can include physical features of the landscape, such as roads, dams, rivers, and fragmented forests, and also less visible barriers, such as the differing ecological, economic, and legal conditions that exist across various political borders.

**Conclusion**

Many scientists consider the present mass extinction event to be the most serious crisis facing the world today. Its diverse consequences ripple out into ecosystems, drawing in and disturbing the lives of countless humans and nonhumans, in many cases affecting the most vulnerable populations first. A great deal of this scientific knowledge is presented in terms of data and statistics. Although we need numerous ways of engaging with and writing about extinctions, one value of a spatial perspective is that it interrupts the hegemony of numbers. In doing so, geographical approaches facilitate an understanding that includes humans within ecosystems, analyzes both connections and interruptions, and is attentive to the cultural, emotional, and political dimensions of extinctions.

*Deborah Bird Rose and Thom van Dooren*

See also Anthropogenic Climate Change; Biodiversity; Biogeography; Conservation; Exotic Species; Game Ranching; Global Environmental Change; Human-Induced Invasion of Species; Keystone Species; Nature-Society Theory

**Further Readings**


Extractive reserves are territories dedicated to environmental protection and the sustainable use of nature resources by traditional populations. Reserves follow a traditional land tenure model based on individual family and communal property rights to common areas, such as forest trails used to extract or harvest nontimber forest products. Although the extractive reserve concept originates in the tropical forests of the Brazilian Amazon, reserves have also been created in aquatic, floodplain, and savanna landscapes throughout Brazil. There are now 50 extractive reserves covering more than 10 million hectares, an area larger than Portugal, and more continue to be created. Despite their growing areal extent, the success of these areas for reconciling conservation and development is still being debated. However, the reserves remain popular with policymakers in part because they address both the land tenure concerns of the local people and the environmental concerns of conservationists. This entry focuses on the forested extractive reserves of Amazonia.

The extractive reserve concept originated with the struggle of Amazonian rubber tappers, harvesters of the latex of the rubber tree (*Hevea brasiliensis*), against the encroachment and deforestation practices of cattle ranchers moving into the Brazilian Amazon in the 1970s and 1980s. Tappers developed a resistance strategy combining nonviolent confrontations with the promotion of standing forests as viable development alternatives and themselves as keepers of valuable forest knowledge. The methods and message of the tappers attracted environmental and human rights groups, who in turn brought international attention to the cause. The rubber tapper movement gained even greater notoriety with the tragic assassination of their internationally known leader, Chico Mendes. Two years later, in 1990, the first extractive reserve, Reserva Extractivista Alto Juruá, was declared.

The creation of extractive reserves generated immediate debate over the economic and conservation viability of both extractivism and organizational units built on the sustainable use of nontimber forest products. Pro-extractive reserve researchers and conservationists found the reserves with great economic and ecological potential for long-term sustainable development, including maintenance of standing forest, biodiversity, and environmental services. Meanwhile, the rubber tappers themselves focused more on the social potential of reserves to generate employment opportunities, preserve subsistence livelihoods, foment local participation in national policy decisions, and support cultural values and local knowledge. Early critics of extractive reserves warned against the idealization of reserves as a panacea for Amazon conservation and cited concerns with the economic sustainability of extractivism (e.g., inelastic supply of the extractive product and low demand), the spatial nature of the targeted resources (extensive nature, low density, and distant from markets), and ecology (potential biotic impoverishment through overharvesting and the deforestation potential of extractivists also pursuing agriculture and animal husbandry). Despite these concerns, extractive reserves gained credence as a means of preserving standing forests and protecting traditional livelihoods.

The complex land tenure arrangement of the reserves, combining public property, community management, and private resource use of designated forest areas, has provided an important refuge for both the forest and the extractivists in the face of continued deforestation and development in the Brazilian Amazon. As road networks, commercial agriculture, and cattle ranching expand into the Amazon basin, extractive reserves increasingly stand out in satellite imagery as
forested islands. However, a closer analysis of the forested reserves reveals some fragmentation taking place as the livelihoods and land use of rubber tappers adjust to new opportunities and constraints.

The traditional livelihoods of rubber tappers included the collection of latex, nuts such as the Brazil nut (*Bortholletia excelsa*), oils such as that from the copaiba tree (*Copaifera* spp.), and even subsistence agriculture and animal husbandry. The collapse of the rubber economy simultaneous to the creation of the reserves forced many tappers to focus on other nontimber forest products and to begin selling agricultural and animal products. Concomitantly, the reserves’ united goals of conservation and social justice encouraged international agencies, governments, and nongovernmental organizations to invest resources and research on improving the economic viability of the reserves through initiatives focused on marketing networks, technical innovation, management, and the search for new and diversified extractive products. These initiatives have provided valuable support to residents seeking to continue the extraction of nontimber forest products in the face of alternative income-earning pursuits such as cattle ranching, logging, and farming.

However, extractivist livelihoods, as in Amazonian forests, are characterized by heterogeneity and dynamism. Thus, while the social value of extractive reserves for rubber tappers and the superior conservation value of reserves in comparison with the expanding cattle ranches that spurred their creation are undisputed, some residents are not practicing the same livelihoods envisioned by reserve proponents. In some reserves, this runs counter to the management plan established for the extractive reserve. This begs the question of who should enforce the management plan: reserve residents or the Brazilian environmental agency. To date, enforcement has been infrequent, raising concerns about the long-term future of reserves. Extractive reserves provide an important opportunity to study the dynamism of Amazonian livelihoods and the challenges to reconciling conservation and development within static organizational units. Perhaps most important, these forested units serve as home and workplace for their residents, even as the reserves become increasingly important to mitigating deforestation and conserving biodiversity in a rapidly developing Amazon basin.

*David S. Salisbury*

See also Biome: Tropical Rain Forest; Conservation; Environmental Protection; Environment and Development; Forest Fragmentation; Indigenous and Community Conserved Areas; Indigenous Environmental Knowledge; Indigenous Environmental Practices; Indigenous Forestry; Indigenous Reserves; Indigenous Water Management; Political Ecology

### Further Readings


### Extreme Geography

Extreme human geography is a critique that celebrates odd juxtapositions; it is not to be confused with extreme physical geography, which has its own distinct cabinet of remarkable curiosities. Extreme geography as a recurring subversive idea in the history of geographic thought manifests as fantasy mapping, counterfactual geographies, science fiction spaces/places, and other imaginative social critiques. As an identifiable critical “tradition” in modern Western geographic thought, its convoluted heritage is cross-disciplinary and can
be traced in fits and spurts from Homer, through Diogenes of Sinope, Saint Francis of Assisi, and Athanasias Kircher, to Guy Debord, Ursula K. Le Guin, and recently a conspiratorial core of radical cartographers at Pennsylvania State University, the founders of *Globehead! Journal of Extreme Geography*.

In March 1994, a graduate student, Nikolas H. “Ni4k,” Huffman described himself in *Globehead!*’s founding editorial as a shameless communist who will say anything, especially for money. Huffman and fellow maverick editors exhibited anachronistic glee while parading *Globehead!*’s outrageous, often incendiary maps, essays, illustrations, photographs, recipes, and classified ads. The editors invited contributors to push the limits of their own geographical imaginations and, in the manner of extreme sports, to brave the boundaries of the geographical and think the unthinkable. *Globehead!* spearheading the extreme geography idea, thus set itself out to be geography’s problem child by “aping the Game” and its counterfeit seriousness and playing up to irony and indignation.

By posturing themselves as extreme geography thinkers and practitioners against disciplined academic solemnity and reveling in ridicule and the ridiculous, *Globehead!* participants exuded hyperradical postmodern attitudes at a time in the history of geographic thought when the bloom was new on the rose of the epoch of postmodern geographies. In spite of the initial enthusiastic endorsements and contributions by several prominent academic geographers, including Peter Gould, Michael Dear, and Yi-fu Tuan, *Globehead!*’s second issue was fatally over the top for an academic geography journal in its time. *Globehead!* published the irreverently illustrated article “Fucking Geography” in its second and last issue, proving that extreme geography is forever born a doomed lark destined to become victim to its own exuberance.

Michael Dear, in his epigraph to the founding issue of *Globehead!* claimed that it was the best idea since postmodernism “came along to confuse everyone.” Yi-fu Tuan wrote that he initially ventured into *Globehead!* because he felt that he belonged there: His own extremism at that time hardly ever touched base with any of the pillars of late-20th-century geography—Marxism, poststructuralism, feminism, and so on—all of which, from his point of view, were by then recognized (and easily labeled) positions and hence mainstream.

Position-less, dysrational, devoid of lasting appeal, and self-destructive, *Globehead!* extreme geography played a brief nihilistic game at Penn State, in odd juxtaposition to its rational surroundings. However, as a recurrent social critique in the history of geographic thought, the absurdist extreme geography idea—much like radical relativism—is easily refuted though never long retired.

David J. Nemeth

See also Critical Human Geography; Gould, Peter

**Further Readings**


**EXURBS**

Exurbia is low-density residential development that occurs at the suburban edge of metropolitan regions. The stereotypical vision of exurban life includes large homes on spacious lots, family life centered on the home, and a desire to escape from the urban ills of crowding, crime, and congestion. However, in contrast to those who see paradise in the guise of exurbia, critics of exurban development contend that such low-density landscapes are simply a form of sprawl where resources are squandered on large lots and long commutes.

Exurban areas are characterized by high commuting flows, low housing density, and rapid population growth. According to the Brookings Institution, about 17 million people, 5.6% of the U.S. population, lived in the exurbs by 2007, compared with 14 million people in 2000. In 2009, the U.S. Census Bureau released a study


called “Population Change in Central and Outlying Counties of Metropolitan Statistical Areas: 2000 to 2007,” which reported that the outlying areas of large cities—those areas beyond the traditional suburbs—grew more than 13% between 2000 and 2007, nearly double the average rate of the closer suburbs.

Why are exurban areas growing so fast? In his 1955 book, The Exurbanites, Spector (1955) argued that exurbanites were motivated by a search for a rural environment. The exurban rural dream is a residential landscape in close proximity to environmental amenities such as open space. Spector’s argument was echoed by Leo Marx (1964), who emphasized the role of the rural ideal in shaping the patterns of U.S. landscape development. And Friedman and Miller (1965) argued that exurbia represented a new type of urban form, created by exurbanites motivated by the desire to enact long-standing beliefs in the efficacy of rural landscapes.

The notion that exurbia represents a new settlement form is a controversial one. Thus, much of exurban research attempts to discover whether exurbia represents a “clean break” from previous settlement patterns or whether exurban development is simply suburban development, albeit in a low-density format.

In an effort to discover whether exurbanites are motivated by a different set of residential preferences, Crump (2003) compared the motivations of suburban and exurban residents and found that there are significant differences between the two groups. For example, exurbanites place a significantly greater value on rural characteristics such as open space, while suburban residents express a greater desire for access to transportation and shopping.

In contrast to these findings, other studies find that exurbanites are really no different from suburban residents. Nelson and Sanchez (1999) found that there is little or no evidence to support the “clean break” hypothesis. Instead, their findings indicate that there is little difference between suburban and exurban residents.

Beyond issues of residential preference, economic factors clearly help fuel exurban development. In particular, the decentralization of employment helps drive exurban growth. Also, the expansion of so-called edge cities, with their large employment nodes and retail centers, is reflective of the growth of employment at the edge of metropolitan regions.

In addition to the decentralization of employment, exurban growth likely reflects the rising cost of housing in an era of income stagnation for many households. Berube, Singer, Wilson, and Frey (2006) found that many exurbanites are trading lower housing costs for longer commute times. This “drive to qualify” allowed exurbanites to purchase homes that they otherwise could not afford in more expensive suburban locations.

Whatever the forces behind exurban growth, the expansion of exurbia has resulted in large swaths of low-density development at the edge of metropolitan regions. The low-density character of exurban development can be characterized as a form of urban sprawl where developers seek out permissive local governments that exercise little restraint over large-lot development.

Although media images of exurbia rely on stereotypical views of “Realtor Mom” and “Patio Man,” the reality of exurbia is more complex. In a notable attempt to unravel the complexity of exurbia, Hayden (2003) developed a three-way classification of exurban settlement patterns; in her view, exurban towns are reluctant suburbs, hot towns, or Valhallas. Reluctant suburbs are rural places inundated by exurban population growth where newcomers often clash with long-term residents. Hot towns are affluent places populated by wealthy telecommuters, and exurban Valhallas are amenity-based exurbs where the wealthy residents use gates to limit public access and to maintain their prized privacy.

The voting patterns of exurbia were the subject of a great deal of interest during the 2008 presidential contest. In the 2004 election, 57% of exurbanites voted Republican. In contrast, Republicans garnered only 42% of inner-suburban votes. Prior to the 2008 election, pundits wondered whether the largely Republican exurbs could offset the gains the Democrats made in the suburbs. This was, apparently, not the case as the initial results indicated that the Democrats had made significant inroads into exurban precincts.

Today, the viability of exurbia is being tested. In 2008, rising fuel prices greatly increased the
cost of commuting, and the widespread use of exotic (and risky) mortgage products such as adjustable rate mortgages caused foreclosure rates to skyrocket in exurban locations such as the Inland Empire of Southern California. Numerous reports of vacant and blighted exurban properties have appeared as the stagnant pools attract mosquitoes, squatters invade the homes, and skateboarders take advantage of the abandoned exurban swimming pools left behind by the newly dispossessed.

Jeff R. Crump

See also Commuting; Housing and Housing Markets; Suburban Land Use; Suburbs and Suburbanization; Urban Spatial Structure; Urban Sprawl

Further Readings


