Barriers and Bridges to the Renewal of Ecosystems and Institutions

Lance H. Gunderson,
C. S. Holling, and
Stephen S. Light
Editors

COLUMBIA UNIVERSITY PRESS
NEW YORK

Columbia University Press
New York Chichester, West Sussex

Library of Congress Cataloging-in-Publication Data
Gunderson, Lance H.
Barriers and bridges to the renewal of ecosystems and institutions /
Lance H. Gunderson, C. S. Holling, and Stephen S. Light.
p. cm.
Includes bibliographical references and index.
ISBN 0-231-10102-3
1. Environmental policy. 2. Environmental policy—Case studies.
3. Social ecology. I. Holling, C. S. II. Light, Stephen S.
GE170.G86 1995
333.95'153—dc20 94-35061
CIP

Casebound editions of Columbia Universtiy Press books are printed on pernament and durable acid-free paper.

Printed in the United States of America
c 10 9 8 7 6 5 4 3 2 1
p 10 9 8 7 6 5 4 3 2 1

10

Sustainable Development as Social Learning: Theoretical Perspectives and Practical Challenges for the Design of a Research Program

Edward A. Parson and William C. Clark

Meeting the challenges of sustainable development (development that meets the needs of the present without compromising the ability of future generations to meet their own needs [World Commission on Environment and Development 1987]) will require substantial advances in our understanding of how natural and social systems interact over long time periods and large spatial scales. Elements of the needed theory of ecological and earth system dynamics are beginning to emerge, as illustrated by many papers in this volume and the growing understanding being generated by research on global environmental change. Substantial progress has also been made in understanding the human side of the sustainable development equation. Most of that progress, however, has been from a static perspective that treats environmental change as either a consequence of, or an exogenous shock to, social systems. The bias of most research and policy programs has been toward institutions, processes, and practices that might maintain or restore some presumptive "equilibrium" with nature (Clark 1989). Generally lacking are theories of social dynamics that can complement the emerging theories of ecosystem dynamics to produce real understanding of the long-term, large-scale interactions of environment and development.

Among those social theories that are dynamic, the most striking common feature is reference to learning. Learning, as a manifestly dynamic process, provides a theoretical counterpoise to social theory based on power and interest, which characteristically yields a static formulation (Adler 1992). The tension between theories of learning and power, and their interpenetration in explanations of social change and stability, permeate the great social and political theories from the classics, through the Enlightenment, and into the present century. Our goal in this chapter is to explore what theories and metaphors of "social learning" might offer efforts to understand the barriers and bridges to sustainable development.[1]

The term *social learning* conceals great diversity. That many researchers describe the phenomena they are examining as "social learning" does not necessarily indicate a common theoretical perspective, disciplinary heritage, or even language. Rather, the contributions employ the language, concepts, and research methods of a half-dozen major disciplines; they focus on individuals, groups, formal organizations, professional communities, or entire societies; they use different definitions of learning, of what it means for learning to be "social," and of theory. The deepest difference is that for some, *social learning* means learning by individuals that takes place in social settings and/or is socially conditioned; for others it means learning by social aggregates.

This chapter is organized as follows. Some philosophical tensions that run through the concepts of learning and of social learning are discussed in the section after the introduction. Theories of individual learning and its social determinants are reviewed in the next section, followed by a section that discusses related work in which the perspective is shifted more strongly to the social level. After a general discussion of relationships between learning in collections and individual learning, the literature on learning in formal organizations is surveyed. The possibly unique situation of learning in science is considered in the next section, followed by reviews of theoretical and empirical studies of learning in politics, policy-making, and international relations. Evolutionary theories in social science and their relationship to learning theories are discussed in the next section. Tentative conclusions and assessments of how the literature of social learning might help to guide empirical research programs on sustainable development are presented in the final section.

Though our main focus is on theory, we have included in the review some applied and empirical work that uses the social learning concept and that shows parallels to this project. The overall approach is inclusive: a work's ostensible concern with social learning was a sufficient, but not a necessary, condition for its inclusion. Consequently, we have cast a broad net, and the treatment of many rich bodies of literature is highly schematic.

The long time scales relevant to sustainable development introduce one tension that we state but do not resolve: On such time scales, longer than many careers and many lives, substantial influence over the development of an issue may be beyond the capacity of any individual agent. Since concern for sustainable development is at least partly normative, however, it makes sense to apply evaluative criteria to decadal-scale social response. Some societies may do better than others in identifying significant risks associated with the interactions of environment and development, in marshaling research and information bearing on the risks, and in managing them. It is our hope that learning about long-term social learning processes may help some agents better understand how to play a constructive role or may suggest ways of structuring institutions, organizations, or negotiations to make effective learning more likely.

Philosophical Roots

Some of the conceptual differences among different students of social learning reflect deeper philosophical tensions that lie within the concept of learning itself, and in social learning in particular. On the confusing nature of learning in general, Polanyi (1966) cites the paradox first raised in the Meno—that to seek new knowledge and recognize that you have found it, you must have somehow already known it. Polanyi's resolution is that all knowledge, in addition to its explicit, conscious part, has a second part that is implicit, or "tacit," and that learning involves attending to different parts, different levels of a question or task in a way that moves bits of knowledge back and forth between the conscious and tacit realms. In anything that we really know, we know more than we can tell. Polanyi wrote only about individual knowledge, drawing on examples of language, such complex skills as music and chess, and scientific discovery; others have used his conscious/tacit dichotomy to refer to knowledge and learning of groups as well.

On social learning in particular, Friedmann (1987) argues that the term embodies tensions that were first articulated in John Dewey's epistemology and politics. Dewey argued that all knowledge comes from the interaction of people with the material environment and that collective learning from such practical experience, guided by the principles of scientific inquiry, could lead society progressively toward a golden age (Dewey 1927). Public decisions are viewed as a series of experiments with the world, which will yield progressive social improvement.

Dewey advanced several views on the difficult question of the validation of such learning and of knowledge in general. In "The Quest for Certainty" he proposed the pragmatic condition that knowledge is validated when it helps an actor settle a problem, granting that its "settledness" may be provisional or temporary. This condition, although reasonable for individual learning and consonant with the "learning-by-doing" approach of Dewey and modern students of innovation, poses problems for public decisions. It presumes that the settledness of an issue is not persistently opaque, as may be the case with certain issues of risk management. And it does not address the question, "Settled for whom?" or the possibility that whether a public issue is settled may be deeply contested.

Dewey later moved to the consensual theory that knowledge is validated by people's opinions (1980). He thus regarded the validation of knowledge, although not its production, as a social enterprise. But whose opinion validates? This view still presumes that no concepts are "essentially contested" and that there is no fundamental conflict between the primacy of public opinion and a scientific, experimental approach to public decision making.

This last tension, between democracy and reason, appears most sharply in "The Public and Its Problems." Here Dewey espoused an idealistic vision of neighborly discourse as the foundation of all political decision, but he also argued that most of the concerns of the modern state are technical matters, best left to the experts to resolve on the basis of the facts. He thus bequeathed to the modern debate on social learning its two most enduring tensions, one positive and one normative: the roles of the individual and of society in the creation and validation of knowledge; and the appropriate relationship between a learning, conversing public and political or expert authority in public decision making. We will see that these run through all the streams of literature surveyed here.

Individual Learning

The process of learning is so central to what it is to be human that any model of the person will address it. Theories of learning are consequently just as diverse as models of the person. Two extreme points for models of learning are marked by the rational-actor and the radical behaviorist models. In their pure forms these two share the characteristics of great cogency and parsimony, and strongly mixed empirical success; they differ in almost every other conceivable way. They embody venerable philosophical disputes on the nature of the mind, and pose strongly differing methodological agendas. As general models of the person, they are so well known as not to need summaries here, except to note that each in its way admits an extremely limited conception of learning.

Rational-actor theory—rooted in economics, political science, and philosophy—is a normative theory that also makes descriptive claims. It is first a theory of choice, but also implies a theory of rational belief, because to choose, individuals must predict the results of their choices. It is in forming and revising estimates of the consequences of choice that the rational actor can learn (Elster 1986).

The rational actor's learning is limited, though, because the actor already knows so much. Only two kinds of learning can occur: Bayesian updating of probability distributions of world states, and in interactive decision situations, reassessments of other agents' interests, available choices, and rationality. More substantial revision of cognitive structures, decision algorithms, or world views is excluded by the presumption of optimality. As for changes in interests or values, the theory does not exclude them, but it has nothing to say about them; it is purely instrumental.[2] The social context of choice and learning is reduced to its informational content—evidence to be used in rational reassessment of the decision situation.

The behaviorist actor's learning is also limited, because the actor knows so little. The methodological focus is strictly on behavior, viewed as a set of stimulus/response connections determined by the past reinforcement regime. The general principle is that behaviors that have been reinforced (rewarded) will occur more frequently in the future. Although the behavior exhibited can be versatile, complexly contingent on environmental conditions, even seemingly creative, "learning" is deemed identical to the observed changes in behavior. Concepts of retained information or image, or internal cognitive structures, are deemed unobservable and superfluous.[3] The social context of choice and learning is reduced to its reinforcement content—a set of socially delivered rewards and punishments.

If the rational actor has a nearly perfect internal model of the world and the behaviorist actor has none, any theory that can accommodate meaningful learning must lie between the two. Actors will hold incomplete, imperfect models of the world, subject to revision in the light of experience. There have been many branches of such theory over the last couple of decades, some growing from the behaviorist root and some from the rational- actor root. All these branches are informed by the empirical results of an increasingly powerful experimental technique. The discussion here briefly reviews a few of the major branches. At the same time, more theorists have looked at the social context of learning. In a somewhat arbitrary division reflecting historical lineages, one body of work stressing social factors in particular learning tasks is discussed here, whereas another that stresses social determinants of broader cognitive development is discussed in the section on codetermined individual and social learning.

One influential body of learning theory is known as Social Learning Theory (SLT).[4] SLT is strongly rooted in behaviorism, but it broadens the range of phenomena that are taken to affect individual behavior. In particular, it grants more standing to social determinants of individual learning.

SLT introduces three factors that determine behavior in addition to directly experienced reinforcement: observation and imitation of the behavior of others; symbolic representation of events and experience through language and other media; and self-generated rewards and punishments. Together these enable people to transform the stimuli impinging on them, and thereby partially to control their own behavior. This intermediate degree of individual autonomy, in Bandura's (1977) words, "neither casts people into the role of powerless objects controlled by environmental forces nor makes them free agents who can become whatever they choose."

The emphases on symbolic systems as reinforcement intermediaries and on self-generated reinforcements both suggest a role for internal representations that puts SLT at some distance from its behaviorist heritage. Different strains of SLT are more or less "representationalist," but the trend has been toward more representationalism (Bandura 1973; Bandura 1977; Aronfreed 1976; Mischel 1968; Kanfer 1971; Wren 1982; Rushton 1982).

A second major line of psychological theory, cognitive dissonance theory, focuses on changes in attitude and belief (Festinger 1957; McGuire 1966). This theory views belief systems as highly interconnected and people as seeking to maintain the coherence of these systems. Inconsistencies among cognitions, or between cognitions and behaviors, cause uncomfortable tension ("dissonance") that people seek to reduce by resolving the inconsistency. The standard experimental demonstration involves inducing people to act contrary to their beliefs and demonstrates that their beliefs then adjust. The less the externally imposed pressure or justification for behaving as they did, the greater the adjustment. The fundamental, highly counterintuitive result is that attitudes can adjust to behavior rather than the reverse, or in Bruner's (1979) words, "People act themselves into a way of believing as readily as they believe themselves into a way of acting."

There has also been development of new theory from the rational-actor pole. The rational-actor model has been softened by economists and students of artificial intelligence who realized that rational actors faced impossible computational and observational burdens. The goal is to relax assumptions of omniscience and reflect people's real perceptual and cognitive limitations while still retaining an evaluative framework.

Evidence of the need for such movement has been of two kinds: (1) empirical evidence of how people choose, learn, and solve problems and (2) attempts to simulate these processes on computers. Researchers in artificial intelligence have found that heuristic structures to filter and interpret information, and to short-cut lengthy optimizing procedures, are essential for computers to address decision problems of complexity that people handle routinely (Tamashiro 1984; Newell and Simon 1972). Studies of people's decision making, learning, and problem solving, on the other hand, show the working of strong ordering principles that are not describable as rational optimization. There are strong, systematic biases in the attribution of causes to other people's behavior (Heider 1958; Schneider et al. 1970; Jervis 1976), in the estimation of probabilities, and in inferences of causality (Tversky and Kahneman 1974; Hogarth and Reder 1986). Choices in uncertain situations depend strongly and systematically on the way the situation is described, or "framed" (Kahneman and Tversky 1981).

Early theoretical models moving away from the strict rational-actor approach were called *bounded-rationality* models (Simon 1955, 1959), stressing human computational limits that prevented optimization. The simplest such models originate in cybernetics and use a model of control

(and learning, to the extent that it occurs) based on negative-feedback loops along a few simple information channels (Bateson 1967). More sophisticated versions explicitly substitute for optimization the search for an acceptable alternative, subject to information, calculation, and other constraints.

A more recent approach focuses on information and its processing, particularly on the need to select, filter, and impose structure on the vast excess of information coming in through the senses. People are viewed not as passive recipients of information but as active shapers of their experience, even at unconscious levels in the process of perception (Steinbruner 1974; Hochberg 1964). They construct powerful but fallible knowledge structures to make inferences about the environment and update them like "naive scientists" (Nisbett and Ross 1980). These structures reduce both information needs and ambiguity in the environment, to make it stable enough to allow people to choose and act.

Modern approaches largely share a focus on the internal cognitive structures used to represent the world. The convergence of these disparate bodies of research is called *cognitive science.* Gardner (1987), in a history of the cognitive science movement, summarized its distinguishing characteristics as follows: a focus on information processing; a reduction in the influence of context, affect, culture, and history; explanation using representations and internal cognitive structures; use of computers as analogies and research tools; and rootedness in old philosophical questions about the nature of mind (Hunt 1989).

Different kinds of mental representations are posited by researchers studying different mental phenomena—perception, categorization, imagery, problem solving, and memory. In one influential body of work, the constructs are called schemas (Minsky 1975; Schank and Abelson 1977; Rumelhart 1980). These are generic concepts stored in memory, collections of usual, representative knowledge. Schemas are not attitudes; they have no affective significance or evaluative component and are purely cognitive. They serve several functions: the selection of what is important from sense data; economical storage of information in memory (which they achieve by reducing redundancy); facilitation of "reasonable" inferences that go beyond the information available; and promotion of the envisioning and carrying out of sequences of actions. Understanding is viewed as a process of matching new information to existing schemas, with some analogic reasoning required, since every new stimulus has some unique features.

There are three types of schemas: scripts, metaphors, and personae

(Abelson 1976; Larson 1985). Scripts are characteristic sequences of events generalized from experience, such as "going to a restaurant."[5] Scripts mediate behavior in two ways: by the selection of a particular script as the best representation of a situation and by the choice of a role in the script. Metaphors are schemas selected as analogic models when there is no literal fit close enough to the present situation. A common example is the representation of complex situations in terms of sports or games (Schon 1979; Miller 1979). Personae are stock characters that permit inferences of someone's likely personality or behavior from their superficial resemblance to a persona (Nisbett and Ross 1980).

The three characteristic inferential biases observed by Tversky and Kahneman (1974)—availability, representativeness, and anchoring—fit well with a schema approach. Availability and representativeness describe efficient but fallible procedures for the selection of a particular schema to apply to a situation: the schema chosen usually is vivid, recent, or memorable (availability) and shows certain stereotypical, superficial similarities to the case at hand (representativeness). Anchoring describes the persistence of schemas once chosen and their resistance to change in the face of contrary information.

The economic storage of information that schemas offer, though, comes at the cost of a loss of flexibility in accommodating new situations or novel combinations. Holland and colleagues (1986) argue that a more general representation of human inference and decision making should be based on systems of mental rules, clustered in hierarchical structures. Rules can be synchronic or diachronic in their model and specific or general in their application. Several rules can be triggered simultaneously by any particular state of the environment; if simultaneously triggered rules conflict, other rules define precedence relations, and rules can gain or lose strength through competition over time. Induction processes include both parameter revision, such as changing rule strengths, and mechanisms to generate plausible new rules.

The list of characteristics that Gardner cited to define cognitive science did not relate explicitly to learning, and different cognitive theories address learning to different degrees. But a simple elaboration of Gardner's list would give the following characterization of cognitive theories of learning: Learning is an experience-driven change in the internal cognitive structures used to represent information. People respond to disparity between their cognitive structures and feedback from their behavior by revising their cognitions.

This description still subsumes much variation. The changing representations can be of very different kinds. Change can be incremental, quantized, or some combination of the two. And change may be principally mediated by motivational factors ("hot control," more connected to behavioral models) or informational ones ("cold control," more connected to rational-actor models; Nisbett and Ross 1980).

Codetermined Individual and Social Learning

Other streams of theory are substantially more social and less individualistic in their orientation than those reviewed earlier. They view social factors as completely dominant, or individual learning and the social environment as interacting so strongly that they are jointly determined. Either view calls for a less individualistic methodological approach.

One line of theory is descended from the sociology and anthropology of the early to mid-twentieth century, which argued that social constructs effectively constrain individual action and thought, and that these constraints are not themselves coherently reducible to individual phenomena (original statement of principles by Durkheim 1938). In its strongest form, as, for example, in the "thought style" shared by members of Fleck's (1979 tr.) "thought community," such a construct "sets the preconditions of any cognition, and it determines what can be counted as a reasonable question and a true or false answer" (Douglas 1986). The most fundamental organizing principles of thought—similarity and difference, classification, and causality—are seen as conditioned by the thought style. And the thought style is invisible to the individuals who participate in it. "The individual within the collective is never, or hardly ever, conscious of the prevailing thought style, which almost always exerts an absolutely compulsive force upon his thinking, and with which it is not possible to be at variance" (Fleck 1979).

For these theorists the focus of inquiry is the role rather than the person, and the norms or rules that constitute a role. The characteristic form of explanation is "functional," in which social forms are explained by the social benefits they bring about.[6] There has been substantial theoretical controversy over the possibility of sustaining coherent, autonomous functional explanations that do not collapse into either individual human intention or evolutionary selection (Dore 1961; Homans 1964). Moreover, functional explanations are static, and hence of limited value as theories of learning.[7] Granovetter (1985) calls such explanations "oversocialized," arguing that to assume rigid social determi-

nation of individual behavior is as limiting as "undersocialized" rational-actor explanations. He advocates a focus on ongoing, changing networks of social relations. Giddens (1984) presents a comprehensive theoretical attempt to reconcile networks of social control and individual human agency.

A more dynamic system of theory that put social factors as dominant causes appeared in the Soviet Union in the 1920s and 1930s, in the work of L. S. Vygotsky and his students (Wertsch 1985a, 1985b, 1991). The primary focus was on cognitive development and its interaction with cultural factors, particularly with tools and systems of signs. Vygotsky contended that culture affects not just the contents of thought, but the structure of its processes.

Wertsch states that three themes form the core of Vygotsky's theory: "A reliance on a genetic or developmental method; the claim that higher processes in the individual have their origin in social processes; and the claim that mental processes can be understood only if we understand the tools and signs that mediate them" (Wertsch 1985)—and that the focus on tools and signs is primary, for these are the socially determined elements that play an essential mediating role in all thought and action, and in the development of thought through instruction.

Vygotsky argued that the development of tools of thought depends on the child's entering into, and subsequently internalizing, a dialogue. The dialogue takes the form of guided participation in tasks slightly beyond their present ability, led by adults or more skilled children (Rogoff 1990; Bruner 1986). Vygotsky defined the zone of proximal development as that set of tasks that a child cannot complete unguided but can complete when guided or prompted with hints, questions, or examples. Vygotsky's contention that all learning takes place in the zone of proximal development has obvious parallels to Polanyi's "tacit knowledge," but on a social plane. Indeed, Rogoff (1990) broadens the view of the dialogue that guides learning to include tacit and nonverbal communication; she also stresses that the learning child is not passive in this process, but actively contributes to the interactive processes that in turn shape her own development.

In this view social causes and active individuals reach a dynamic equilibrium in which social/cultural forms and the cognitions of individuals are codetermined. Cole (1985) has argued that culture and cognition create each other in the zone of proximal development. Wertsch (1991) argues that individual and collective factors are so tightly interwoven that efforts to separate them and look at reciprocal effects in a research program are bound to be fruitless. He argues for a different level of analysis, in which "mediated action" is the primary unit of investigation. Such a research program would pose daunting methodological problems and would likely be cross-cultural.[8]

Organizational Learning

The next three sections discuss learning in various social aggregates: task-oriented groups, formal organizations, and professional communities. As we stated earlier, using the term *learning* to apply to a collection implies one of two forms of relationship between individual learning and changes taking place in the aggregate, which we call decomposition and analogy.

Decomposition treats group learning as the sum of learning by the group's constituent individuals (Udehn 1987; Elster 1989). It is not restricted to the obviously trivial case in which the individuals' group membership does not affect what they learn or how they learn it. What each individual learns may be complexly contingent on the choices and learning of other group members (e.g., in the pursuit of high-level coordinated performance by a group such as a basketball team or a string quartet). Or the means of individual learning might be through activities that depend on the participation of other group members, such as discourse, imitation, or shared activity (Bandura 1977; Argyris and Schon 1978; Habermas 1979).

Analogy treats group learning as autonomous, determined by group-level causal processes that correspond to the processes shaping individual learning. This view may simply represent a methodological convenience; even if all learning is reducible in principle to individuals, the most fruitful way to study groups may involve observations and theoretical constructs at the group level. The value of studying chemistry is not diminished by its being in principle reducible to physics. One may apply analogies between individual concepts—such as perception, reinforcement, memory, cognitive dissonance, or schematic change—and changes in group routines, stories, or behavior, without actually believing that the group sees, thinks, or remembers.

Alternatively, one may reason by analogy to individual phenomena and believe that the collective phenomena are autonomous, truly irreducible to individuals. This position originates in Durkheim's insistence that social facts be explained by social facts and his denunciation of dipping into the psychological level for explanations (Durkheim 1938).

Studies of learning in formal organizations include some instances of both decomposition and analogy, and some studies that combine the two approaches.[9] Although one body of theory, the neoclassical theory of the firm, analyzes organizations by analogy to the individual rational-actor model, the more frequent approach is to study organization-level variables, such as routines and procedures, with analogies to individual reinforcement or cognitive structures.

March and Olson (1976) presented a general model of interaction between an organization and its environment, based on a four-element feedback loop. The cognitions and preferences of individuals within the organization determine individual behavior, which determines organizational choices, which determine environmental consequences, which determine the preferences and cognitions of individuals in the organization. They describe this pattern as a complete learning cycle.

Each of these four causal connections can be broken, though, yielding a particular kind of incomplete learning cycle in each case. When individuals in an organization are not able to act out their preferences, the resultant incomplete learning is called *role-constrained learning.* When individual actions are not translated into organizational actions, the result is *audience learning.* When the causal relations between organizational choice and outcomes are not well understood, the result is *superstitious learning.* And when it is not clear to individuals what the consequences of organizational action were, the result is *learning under ambiguity.* Hedberg (1981) points out that when learning cycles are incomplete, direct empirical falsification does not occur and that consequently, organizations can persist in incorrect belief or ineffective action.

March and Olson's treatment is general and explicitly decomposes organizational learning into learning by the constituent individuals. A focus on organizational routines is also common (Cyert and March 1963). Heiner (1983), for example, argues that simple decision rules or routines are the appropriate response to an uncertain environment whose complexity exceeds the analytic capability of the agent. The larger the disparity is between environmental complexity and the agent's cognitive ability, the simpler and more predictable the decision rules are.

In a recent survey article, Levitt and March (1988) stress reliance on routines in their three basic observations on organizational learning: that behavior in organizations is based on routines, chosen more on a basis of legitimacy or appropriateness for the situation at hand than of outcome calculation; that organizational actions are determined

strongly by interpretations of history and past outcomes and only adjusted incrementally in response to feedback (Lindblom 1959; Steinbruner 1974); and that organizations are oriented to targets, with success or failure assessed by the relation between observed outcomes and aspiration values (Simon 1955; Siegel 1957). Learning is more typically a response to shortfall or scarcity than it is to success, although affluent organizations may search out new opportunities (Hedberg 1981; Lewin and Wolf 1975).

Organizations learn by encoding inferences from history into routines that guide their behavior. In the simplest terms, a strict analogy to the behaviorist model of individuals applies; routines that are associated with successful attainment of targets will tend to be repeated (Cyert and March 1963). If the organization's environment and tasks are sufficiently stable, such learning will lead to long-term trends of improved performance. The clearest observed example of such learning is the "learning by doing" effect, through which production costs in a wide variety of goods and services decline with cumulative experience (Levitt and March 1988). A recent study has shown, though, that the learning effect of past production decays markedly with time (Argote et al. 1990).

Organizations and institutions can also be viewed from a rational-actor perspective as systems of rules and relations that rational agents construct to reduce transactions costs and defend themselves against fraud and exploitation (Williamson 1975, 1985). Organizational change can also be studied as rational response by organization members to a changing environment. In his study of institutional change, North studies the interaction of institutional constraints, organizational routines, and incentives of individuals within organizations to acquire new skills and information. Drawing on Polanyi (1966), he argues that organizational routines, like individual knowledge, embody both conscious, transmittable knowledge and tacit knowledge reinforced by practice and interaction (North 1990).

A more "cognitive" approach would recognize that organizations exert some control over what stimuli to notice; that the assessment of success and failure is problematic and related to internal organizational conflict; that the collective memory in which routines are recorded is imperfect; and that organizations share stories, assumptions, beliefs, and myths that guide action and give meaning to experience but retain important elements of ambiguity (Hedberg 1981; Levitt and March 1988).

Modeling organizational learning as a response to outcomes that are finally interpreted as either success or failure implies that learning is

limited to means of achieving specified organizational goals. Argyris and Schon (1978) consider the deeper process of articulating and reconceptualizing organizational goals and perceived causal relationships, which they call "double-loop learning." Their model focuses on the difference between espoused theories and the "theories in use" that actually guide individuals' behavior in organizational contexts. Like Beer's (1972) "meta-logic," theories in use are typically inaccessible to the individuals who are guided by them. Changes in theories in use are inhibited by cognitive and behavioral feedback loops, which, among other functions, suppress recognition of the present theory. They propose a dialectical process of articulating, and thereby transforming, the unspeakable assumptions that guide organizational action. Though often treated as such, their theory is more than a simple distinction between learning means and learning ends; its connection to critical theories of social transformation (Habermas 1976) is direct.

Like individuals in Social Learning Theory, organizations do not just learn from direct experience; they can learn from observing others. The experience of other organizations can be acquired through the movement or imitation of technologies or procedures, contacts with or movement of personnel or consultants, or professional associations or public media. Organizations, like individuals, have some power to select or control their environments, and consequently to control their learning context. The movement of innovations between organizations has been modeled as broadcast and contagion processes,[10] sometimes with different degrees of match between the innovation and the organization being the most significant factor in explaining adoption (Mansfield 1968; Kay 1979).

Theories of organizational learning can be mapped onto the behavioral/cognitive/rational distinction presented earlier for individual learning. But theories of organizational learning also reflect differences in the fundamental metaphors used to characterize organizations. Morgan (1986) delineates different bodies of work that treat organizations as machines, organisms, cultures, political systems, "psychic prisons," and instruments of domination. Each of these metaphors implies a different view of the nature and determinants of organizational change, and of what it means for an organization to learn.

Although the goal of organizational learning theory has been to improve organizational performance, in 1981 Hedberg concluded that little progress had been made in the prior 20 years and that current knowledge did not permit much specific guidance. He offered a few general

practitioner's suggestions: Promote experimentation by reducing penalties for failure, build in regular shocks to routines through such measures as time-limited management contracts, and don't filter the information reaching senior management too much (Hedberg 1981).

Learning in Science

Science may represent a special case of collective learning. Many social writers have granted a special place to scientific knowledge, believing that the superior verification of scientific method made the resultant knowledge more reliable and less liable to social explanation than other kinds of knowledge. Mulkay (1979) describes the "classical sociological view of science" as holding that scientific institutions and the social use of scientific knowledge were fair ground for sociological analysis, but that the conceptual content of scientific knowledge was not. The perception that science succeeds uniquely in generating new, reliable social knowledge lies behind the once common call for scientific principles to be applied to the solution of our social problems.[11] Contrary views have focused on the usefulness and legitimacy of transferring scientific thinking to the realm of public decision making, rather than on the special epistemological status of scientific knowledge (Lindblom 1959; Lindblom and Cohen 1979; Braybrooke and Lindblom 1970).

But this special status has increasingly come under question. It depends upon several presumed characteristics of scientific method: the generation of falsifiable hypotheses and their attempted empirical refutation; observable facts that are unproblematic and uncontaminated by theory; and gradual, cumulative progress (Popper 1959, 1972). But these have all been cast increasingly in doubt by new theoretical constructions of science and by observations of what scientists actually do.

Kuhn (1970) was the first to argue for shifting focus from particular theories to those larger macrotheories that temporarily bound particular scientific specialties, defining promising directions for research questions and criteria for correct answers. Kuhn argued that these macrotheories (which he called paradigms) are determined consensually and tend to persist even in the face of substantial empirical counterevidence. Lakatos (1970) refined the characterization of macrotheories (which he called *research programs*), distinguishing a hard center or "negative heuristic" of fundamental assumptions that are temporarily immune from criticism, from a softer periphery, or "positive heuristic," which provided general guidelines for the generation of particular testable hy-

potheses. Contrary to Kuhn, he argued that more than one research program can be active at one time and that rejection is not sudden but a gradual decay as other programs show more empirical success. Laudan (1977) proposed that macrotheories (which he called research traditions) can evolve over time, because although a core set of propositions is sacrosanct at any time, that set can change over time. He also argued that theories are tested by the balance between the empirical problems they solve and the conceptual problems and anomalies they generate.

These different views of how science works have increasingly been investigated through detailed historical study. Donovan and colleagues (1988) present sixteen detailed case studies of scientific change, examining hypotheses on the role of guiding assumptions and anomalies, and the character of innovations and revolutions.

Mulkay (1979) argues that the newer view, which regards science's epistemological status as more fallible, makes sociological study of science's conceptual content legitimate. Over the past decade there have been increasingly detailed sociological studies of the collective creation of scientific knowledge. Fleck's 1935 study of syphilis (1979 tr.), which argued the social determination of all cognition, was a remarkable early exception to the former deference. Recently, there have been many studies of routine operation of scientific laboratories and of scientific controversies (e.g., Latour and Woolgar 1986; Gilbert and Mulkay 1984).

Some studies have persuasively demonstrated social factors determining the acceptance of particular scientific theories or evidence,[12] but many others have yielded weak or ambiguous results. Laudan (1977) argues that the pursuit of social explanations of scientific knowledge has been too expansive. He accepts that the territory of scientific ideas can reasonably be partitioned into those parts determined rationally and those determined socially, but he contends that sociologists have too readily accepted the narrow, classical definition of rational explanation and thus have claimed an overly large remainder for themselves. He argues for a "history-of-ideas" approach to the study of science, following a methodological heuristic of first trying to account for scientific ideas on rational grounds of successful problem solving, coherence, and consistency with other bodies of theory, and only seeking social explanations when the results of this attempt are weak. The record of social studies of science would seem to support Laudan's agenda. It suggests that social factors play the largest role when a body of work is new, is incomplete, and has limited or unclear connections to more mature

lines of inquiry; as a field matures, the rational pressures become more tightly constraining and the room for social determination of concepts becomes more limited.

Learning in Policy-making

A second special case of collective learning, one of particular relevance to the study of sustainable development, is the learning that goes into policy-making, policy change, and international relations. In contrast to the classical deference accorded to scientific learning, the traditional view of policy change is that it is better explained by power, interests, and coalition alignments than by learning.

The postwar rise of the disciplines of policy analysis and operations research has led to a contrary and excessive view of the role of technical and scientific knowledge, and learning, in policy-making. The following caricature captures the flavor: *Political leaders, in doing their job of making discrete policy decisions among well-defined alternatives,*[13] *realized that the increasing scientific and technical content of the decisions they were required to make called for outside advice. They turned to a professional cadre of analysts, who presented the relevant scientific and technical information. The decision makers, now knowing the relevant consequences of whatever decision they might make, then performed the appropriate balancing of values and interests to arrive at a decision.* (Note that the analysts operated only in a technical arena, and their only audience was the decision maker.)

This caricature is rooted in the reality of some specific policy and operational problems, many concerned with efficient military operations, that were solved in the early days of operations research. Were such separation of the technical and the political possible, the tension between democracy and reason described earlier in the discussion of Dewey would not arise. But for most policy and analysis, it is indeed a caricature. Even when policy makers know what technical or instrumental questions they want to ask, they are often ones whose answers science does not yet (or cannot) know (Weinberg 1972; Häfele 1974). And in an ambiguous, messy decision environment where goals are multiple or contested, they may well not know what questions they want to ask. In this environment, it is not surprising that so much policy analysis is bad, not listened to, or used for nonsubstantive purposes (Clark and Majone 1985; Sabatier and Jenkins-Smith 1988; Dunn 1980; Webber 1983).

But even with a more realistic view of the relationship between policy makers and analysts, there is much opportunity for learning in this system. One key result is that policy analyses often serve a longer-term "enlightenment function" (particularly relevant to a decadal time view), even if they have little immediate effect on particular government decisions (Weiss 1977, 1982). Clark and Majone (1985) point out that evaluating policy analysis using standards drawn from either pure science or pure democracy is liable to condemn too frequently, and they suggest a set of standards that mixes output, input, and procedural criteria. Increasing recognition of the role of deliberation and of the similarity of deliberative processes in policy and in the sciences has led several authors to call on analysts to acknowledge their role as debaters, persuaders, and rhetoricians and to jump into the fray (Majone 1989; Sabatier and Jenkins-Smith 1988; McCloskey 1985). As Rouse (1987) argues, even in the natural sciences there is a substantial collective, procedural element in determining who is right; the rightness of your view is determined by your ability, using the tools and results at your disposal, to persuade others of it. Majone (1980) extends the analogy between policy and science by arguing that policy change follows a process similar to Lakatosian competition among research programs in science.

An increasing number of case studies in policy change have focused on learning as a key and underaddressed issue (Keohane and Nye 1987; Breslauer and Tetlock 1991). The most common focus is on the learning and cognitive processes of top individual decision makers. Some studies apply models of learning from psychological research; others apply their own definitions and models. Jervis (1976) systematically studies the application of theories of individual cognition and perception, and related biases, to international relations. Larson (1985) applies psychological theories of attitude change to the changing views of four senior American officials at the origin of the cold war. Etheredge (1985) examines senior officials' decision making in recurring similar foreign-policy incidents and concludes that they do learn, but only under conditions of crisis. Richard Rose (1991) addresses the special case of "lesson drawing" across nations by senior governmental officials. His emphasis on the essentially political character of the officials' choice to seek rationally defensible "lessons" and to deploy them in the service of particular agendas is a valuable bridge between the political and scientific poles of the learning debate (see also Lee 1993; Majone 1991). Ernst Haas (1990) articulates two models of nonlearning adaptation and one of learning and applies them to a series of case studies of international organizations.

Steinbruner (1974) describes three characteristic biases, in cognitive processes in organizational decision making: grooved thinking, uncommitted thinking, and theoretical thinking. He asserts that these restrict learning and so characterizes organizational learning under a cognitive model as "constrained learning," but he provides no details of the structure or dynamics of learning under such a model.

Peter Haas (1990) takes more of a group focus in his study of the influence on Mediterranean pollution control agreements of a so-called community. He argues that this international group of like-minded officials, mostly scientists, reached a common understanding of the issue relatively early and then exploited their monopoly on relevant scientific knowledge and their positions of control in specialized government agencies to push their governments into strong international agreements.

Heclo's (1974) study of social policy-making in Britain and Sweden finds an important element of social learning among officials who develop policies. Hall (1990), building on Heclo's work, articulates a model of policy learning in the context of the theory of the state. Three central features of the model are that policy is strongly influenced by recent policy,[14] that a key role in advancing policy learning is played by experts in government or by advisers to it, and that states have substantial ability to act autonomously from societal pressures. Bennett (1990), in a study of learning among senior Soviet foreign-policy decision makers from 1973 to 1983, formulates a series of learning hypotheses on the individual, group, and governmental levels.

Two studies of long-term learning by governments are of particular relevance to a study of learning in sustainable development. Cooper (1989) studied the long-term development of knowledge, cooperation, and institutional capacity for combating international movement of diseases in the nineteenth and early twentieth centuries. Hall (1989) studied the development and movement of Keynesian ideas and policies in eight countries. He delineates three approaches to studying the political influence of economic ideas: economist-centered approaches, which turn on the ideas' intellectual acceptability to economists and presume that economists dominate economic policy-making[15]; state-centered approaches, which emphasize the institutional structure of the state apparatus and its recent experience with similar policies; and coalition-centered approaches in which a new idea's success turns on a politician's ability to construct novel winning coalitions around it. Hirschmann (1989), in his "comment" on Hall, stresses the intellectual component of even a coalition-centered approach: "Prior to Keynes, there was no

respectable theoretical position between centralized planning, on the one hand, and, on the other, the traditional laissez-faire policies, with their denial of any governmental responsibility for economic stability and growth."

Hall concludes that ideas do affect policy, but the manner and extent of that influence depend on their economic, administrative, and political viability. He finds four central factors in determining the adoption of Keynesianism: orientation of the governing party, permeability of the civil service, concentration of power over macroeconomic management, and the power of the central bank over policy-making. (Central bankers did not like Keynesianism.)

In summary, these studies show two distinct approaches: studying learning among a few top political decision makers and studying it in the somewhat larger community of policy experts on a particular issue. With the former approach, the focus is clearly on individuals. Consequently, the large body of work on individual learning, attitude formation, and cognitive change applies directly. With the latter approach, the focus is necessarily on a larger and only roughly bounded group. Consequently, as in the preceding discussion of organizational learning, the relevant learning theory—theory of how the group learns—is not well developed. Several authors limit their comments on group learning to the observation that individual learning is necessary but not sufficient for group learning, for routines and conflict can prevent individual learning from being expressed in group activity.[16] The most frequent response is not to look inside the policy community but to ask how their learning affects policy change. On this question the theorizing is largely ad hoc, amounting to the assertion that learning by policy elites does affect policy change; change is not completely reducible to contending political interests and bureaucratic inertia. The theoretical tools for distinguishing the effects of learning from other forces of change, though, are still at an early stage of development.

Evolution and Learning: Complementary Models?

Aside from learning, evolution is the other manifestly dynamic process often employed as a model for change in social systems. An evolutionary approach can cut across the various categories of learning studies we have considered thus far. Evolutionary models have been applied to learning processes at the individual and organizational levels, in science

and in policy-making, and to other processes of social change. This section discusses the biological basis of evolutionary models, their application to learning and other processes of social change, and the relationships between learning and evolution as theoretical models for social dynamics.

The basis of an evolutionary model is the selection among invariant forms leading to differential survival. In biology evolutionary models are based on stably propagating species defined by reproduction within but not across the species boundary; on variation among the individuals comprising a species, at least some of which is heritable; and on differential reproductive success of individuals leading to shifts over time in the genetic and somatic characteristics of the species.

The differential reproductive success of individuals is mediated by many factors, some of which are fixed (some variations are quickly lethal to the individual carrying them); some of which are contingent on particular characteristics of the environment the species presently occupies; and some of which are socially contingent, depending on the distribution of related characteristics or behaviors among other members of the same population. Overall adaptivity or fitness is not absolute, but contingent on many factors.

The shift in species characteristics over time and in response to changes in environment sometimes brings about the creation of new species. It is widely agreed that speciation requires that a population undergo substantial change during a period of effective physical reproductive isolation, which proceeds so far as to create effective biological reproductive isolation if and when the separated populations come back into contact.

The primary locus of selection, though, is in dispute. The standard view is that selection is among individuals, through processes of reproduction and death. However, one can construct scenarios under which marginal populations diverge quickly enough that internally stable, reproductively separate groups come into competition for the same environmental resources, with selection thereby operating among species through processes of speciation and extinction (Eldredge 1985; Curtis and Barnes 1989).

This basic structure of evolutionary thought—a mechanism of novelty generation and one of selective retention—has provided a richly provocative analogy, and sometimes formal models, for students of learning at the individual and group levels. The richest approach has been not to argue that learning shows particular characteristics because

of adaptive advantage,[17] but to apply evolutionary arguments to nonbiological phenomena that show the same processes of (1) generally stable propagation, (2) some means of generating innovation and variation within the basic regime of stability, and (3) a separate mechanism of selection from among variants, such that some are more likely to survive and propagate than others.

Applying evolutionary thinking to ideas and cognitions themselves, Campbell (1960, 1974) has argued that human learning is at root an evolutionary process. Bateson (1972) proposes an ecological approach to broad classes of human cognition and behaviors. Popper's (1972) view of falsification and progress in science is expressed in explicitly evolutionary terms. Toulmin (1972) presented an evolutionary view of concepts in society at large, not just science, emphasizing that concepts that are dominated are not necessarily eliminated; like marginalized subspecies, they can remain available to exploit future opportunities.

At the organizational level, an evolutionary perspective can be particularly fruitful in view of the competitive relations among economic organizations. Nelson and Winter (1982) argue that competition among firms applies selection pressure to the processes of routine generation and innovation that operate within firms. Since the 1970s the emerging field of organizational ecology has sought to apply ecological and evolutionary principles to broader classes of organizations. Hannan and Freeman (1977, 1984) applied theoretical constructs from population ecology to organizations. They distinguish two forms of organizational change: learning from the environment and adapting to change in the environment. Organizations with too much inertia to change fast enough relative to change in the environment will be removed by selection pressure. The more recent literature has focused on the levels-of-analysis problem, arguing that the population is not the only or most useful level at which to study organizations, particularly the origin and diversification of organizations.[18] Young (1988) provides a forceful critique of biological analogies in the study of organizational change, arguing that such fundamental concepts as species, niche, and death are defined ambiguously or circularly.

In their work on mathematical modeling of cultural evolution, Cavalli-Sforza and Feldman (1981) argue that cultural artifacts, or "memes,"[19] can be modeled by the same processes as biological evolution, with two possible exceptions. First, innovations, unlike biological mutations, are typically created with a purpose or in response to a perceived problem, so the determinants of their adaptiveness may be structured differently, and the probability of innovations is higher than for biological mutations. Second, cultural selection may have complex interactions of either direction with Darwinian selection. Although most cultural patterns likely have minimal direct Darwinian effects, some clearly have negative effects (they cite ritual mutilations, particularly female circumcision, and risky sports), and some have clearly positive results (the adoption of agriculture, with resultant increases in local carrying capacity). Boyd and Richerson (1985) develop a "dual inheritance" model, in which formal Darwinian selection is applied independently to genetic and cultural inheritance.

In less formal treatments, Adler (1991) applied evolutionary concepts to government learning, explaining two major innovations in American foreign policy through a process of idea generation, domestic selection, and international diffusion. Schmid (1987) and Eder (1987) take an evolutionary approach to the explanation of social systems, norms, and rules. At the social level of analysis, theories focusing on the cognitive evolution of politically salient ideas have tended to become linked with normative discussions of progress in human affairs.[20]

An evolutionary model cannot be equivalent to a learning model at the same level of analysis. Evolution is based on selection among invariant forms through birth and death processes, whereas learning implies that the learning system survives its changes. Moreover, as Elster (1979) argues, evolutionary systems are restricted to changes through "local hill- climbing," selection based on the immediate adaptive advantage of each local increment, while human learners can proceed obliquely, take one step back for two forward, pause to reframe a problem, or otherwise circumvent the demands of continuous local improvement in pursuit of a superior global goal.

But an evolutionary model may be equivalent to a learning model at a different level of analysis. Selection among organizational routines may represent learning by the organization (Nelson and Winter 1982). Selection among particular organizations by birth and death may represent learning by the system in which the organizations operate. As Campbell (1974) and Popper (1972) have argued in different contexts, individual learning may be an evolutionary process at the level of cognitions, through which "our hypotheses die in our stead."

If evolutionary concepts are to be used in studying social learning, the selection must operate on the things learned—the concepts, names, slogans, images, acts, technologies, forms of social organization, political movements, opinions, attitudes, norms, and skills that people learn,

individually or collectively. The evolutionary concepts of selection, ecological community, and species suggest many provocative analogies, hypotheses, and questions for the study of social learning. To determine how useful they are requires a careful look at the applicability of each of the constituent concepts: origination of variation, propagation, selection, and speciation and species stability.

First we need a description of origination. Although most would agree that particular new ideas are generated by individuals, there is more dispute over the relative effects of individual creativity and social ordering on the total set of new ideas generated. Social or cultural factors may interact with individual creativity to determine the range of novel concepts a member of a particular society is capable of inventing. Thereafter some social factors are likely involved in determining the intelligibility, sense of importance, and ease of communicating a new idea or image, and consequently its initial viability.

Next we need a mechanism of propagation. Different forms of propagation apply to different classes of things learned. Some things can be replicated cheaply and accurately by print or electronic media, others (e.g., behaviors to be learned by modeling) require more detailed information, and still others (e.g., skills and craft knowledge, forms of argument, persuasion to a new value) may resist explicit encoding completely and require direct communication from another person, guided participation, or practice.[21]

Models of information propagation alone may not be adequate to capture the full range of learning processes, though. In addition to the issue of tacit knowledge, there may be things learned that, contrary to the cognitive science model, bundle evaluation or affect with information.[22] The particular content of information and affect may change as a word or slogan is propagated and used, through the effects of both the channels of communication through which it passes and the people who hear and use it.

Third, there must be mechanisms to introduce variation in the propagation. Although physical processes of replicating information have low error rates, people's inclination to interpret and modify what they receive in view of their own experience, values, information, and cognitive processes will create pressure toward variability. One needs only to have played the party game in which a story is passed around the room in whispers, finally returning to its originator transformed beyond recognition, to realize how rapidly an idea can be transformed in interpersonal transmission.

Additional variation can also arise from the movement of ideas through various communication channels and to different audiences.[23] Certain kinds of ideas, such as technical concepts of specialized disciplines, may only be intelligible to a few people. When such ideas move through broader circles, they may be altered in ways that reflect the characteristics of the new, broader audience or of the channels of communication through which they pass (e.g., new scientific results reported on television or used in congressional debate). There is also something in the transmission of ideas akin to sexual reproduction in biology, with its mixing of existing genetic material to create new variations. Ideas, norms, and images come together in inter- or intrapersonal process—in introspection or in discourse—and generate new ones.

If there is to be a strong analogy to biological evolution, though, these forces for variability must be countered by forces to stabilize related collections of ideas. If the intrapersonal processes determining what is intelligible or worthy of consideration are not sufficient to effect this stabilization, then it must come from some interpersonal process—from discourse, from comparison to other cognitions and views held, or from argument and evaluation in view of a common shared body of knowledge.

This stabilizing force could operate as selection pressure on individual ideas to reinforce related systems of ideas. Bodies of ideas could be mutually reinforcing by being intelligible in terms of each other, reciprocally confirming, suggesting questions or actions whose results will further reinforce the collection. Disciplinary boundaries perform these functions in science, as do the various forms of macrotheory within disciplines (Lakatos 1970; Kuhn 1970; Laudan 1977). In political and social discourse, corresponding bodies of thought might be called ideologies or national characters (Geuss 1981; Bateson 1942a). The analogy to selection forces preserving the stability of species is evident.

There is abundant anecdotal evidence for the notion of bodies of related ideas defending their integrity and borders through repression of deviance and sniping at those who enter no-man's-land. This would parallel the argument that the relative paucity of smooth, large transitions in the fossil record shows the middle ground to be highly unstable. If you must change, you survive by evolving quickly across the boundary to find a new stable niche beyond. The often lamented difficulty of establishing and maintaining interdisciplinary scholarly careers and academic programs would fit this model. In this case the niches for which groups are competing are university appointments and research grants.

The competition will be intense. In the realm of political ideas, Chomsky's (1989) work on propaganda argues for the existence of social mechanisms that delimit the range of acceptable political debate in a society.

The adaptation of bodies of ideas, like adaptation of species, is relative. They compete not against objective criteria of value but against each other in view of the opportunities that the environment presently offers. This lends a provisional character to any presently ascendant ideology or thought system.[24]

Moreover, a great breadth of survival opportunities for idea systems will exist in a reasonably pluralistic society. The evolutionary analogy suggests that small populations of marginal ideas, like Toulmin's temporarily dominated concepts, may survive through long periods when the mainstream environment is unfavorable, flourishing anew when environmental conditions change in their favor. The persistence of neofascist movements among alienated young men in industrial countries and the rapid reappearance of such movements (and of ancient ethnic hatreds) in Eastern Europe and the former Soviet Union may be examples of this phenomenon. The analogy also suggests that there may exist many such small, marginalized subpopulations of ideas, collectively representing an important reservoir of cognitive variability that can increase society's resilience to extreme environmental change.

Conclusions: Guiding an Applied Research Program

The purpose of this chapter has been to gather together the major strands of theory related to social learning, with a view to complementing a research program of empirical case studies in the management of sustainable development. The best-developed theories of learning are clearly at the level of individual learners, and to a lesser extent at the level of small, face-to-face groups. Indeed, some writers limit their definition of social learning to processes that occur in such groups.[25] The studies reviewed in this essay, however, suggest that learning can be a meaningful concept to apply to processes that occur at all levels of society, from the individual to the international. A progressive research program on the role of learning in sustainable development must therefore address two sets of questions, one largely definitional and methodological, the other largely empirical.

On the definitional side, those studies are likely to be most useful that clearly specify their approach to four key questions: Who (or what) learns? What kinds of things are learned? What counts as learning? Why bother asking? These simple questions can be used in categorizing works in social learning. They concern what each body of work defines learning to be, what level of analysis they focus on, and how they decide whether or not learning has occurred. These questions are locators; they provide four dimensions on which to map the territory of theories of social learning.

Answers to these questions, perhaps tentative ones, define a research program. Once a focus is defined, a second class of empirical questions arises that the research program aspires to answer. What specific things did they learn? How did they learn them? Under what conditions and why? Did it make a difference? And how could the effectiveness of learning have been increased?

Who or what learns? Societies, governments, and organizations, or just individuals? Our basic conceptions of learning are formed from observing individuals, but collections of people large and small also exhibit clear changes in task performance, coordination, complexity of communication, and goals that look very like individual learning. Any writer who uses the term *learning* to apply to a collection speaks to this common perception, implying some form of relationship between what the collection is doing and individual learning. This implied relationship is typically of one of two kinds, which in our discussion of organizational learning we called decomposition and analogy. With specific reference to studies of sustainable development, at least five different groups of learners seem relevant: senior policy decision makers, scientific communities, industrial organizations making production and technology decisions, other nongovernmental organizations, and—to the extent that consumer market forces and broad political opinions and images condition or constrain elite decision making—the citizenry at large. The role of the mass media, as an agent of learning that may itself learn, is particularly problematic.

What kinds of things are learned? Individuals learn a vast array of different classes of things: behaviors, facts, concepts, words, skills, desires, opinions, attitudes, and values. There may also be a logical hierarchy of categories of things learned, analogous to the difference between improvement in one task and improvement in the rate of learning new, "similar" tasks, which depends on the second-order task of recognizing a context of "similarity."[26] Collections of people learn all the

same sets of things, and such essentially collective things as languages, technologies, slogans, images, forms of social organization, and norms. It is reasonable to presume that different classes of things are learned in different ways. A useful learning theory must be clear about its breadth of application. Useful theory will also treat carefully the relation between "who learns" and "what is learned." Particular things learned will move differently through different communities. Some are only intelligible to particular subpopulations; others may rely on a prior favorable disposition. When taking a macroview of whole societies, particular learners (whether individuals or organizations) may become too fine to resolve, and the appropriate methodological focus may be strictly on the things learned—the ideas, facts, behaviors, and norms—and on their origination, propagation, and growth or decay. In studies of sustainable development, some of the most important things learned seem likely to include scientific facts and models, "policy theories," technologies, preferences, behavioral norms, images and names, or broad world views and conceptions of society.

What counts as learning? Any change in the phenomena studied, or only certain changes? If only certain changes, then what criteria distinguish learning change from nonlearning change? Short of calling any change learning, one might say that only change in response to identifiable stimuli or information is learning; or more restrictively still, that the change must follow rationally (defined somehow) from the stimulus. Some researchers propose narrower definitions still: increasing cognitive complexity, or increased effectiveness at attaining given ends. Others argue that definitionally, any cognitive change should be treated as learning, and questions of effectiveness or progress should emerge from the subsequent research (Nye 1987; Bennett 1990; Breslauer 1987). We have followed the latter advice in this review, since it is the more inclusive approach. Our findings lead us to believe that it will also be a more productive stance for the study of learning in sustainable development.

Why bother asking? Most students of social learning would answer, "In order to understand and explain the societal phenomena we observe." In this view a theory of social learning, like any theory, would be a set of general, causal statements whose purpose is to guide inquiry and explain observed phenomena. Others, working in the tradition of critical theory, take a more activist view of social theory. They would answer that we ask about social learning in order to bring about personal liberation and social transformation. In this view a theory is a set

of propositions to guide communication between people, whose purpose is to bring about these transformations and which is validated by (1) its acceptance by the agent addressed in a noncoercive situation and (2) its efficacy in bringing about the desired transformations (Geuss 1981; Habermas 1979). Although critical theory does not address itself to learning as such, there are obvious points of contact with some of the strains of literature surveyed here—to individual Social Learning Theory as employed in therapeutic practice (Stuart 1989) and to some strains of the organization development literature (Argyris and Schon 1978; Friedmann 1987).

It is difficult to incorporate the critical perspective into a generally naturalistic inquiry into the characteristics and determinants of social learning, but this view cannot be ignored.[27] An inquiry directed to social intervention, as arguably any project in "policy research" is, may lie closer to the agenda of critical theorists than to that of positive theorists. In policy research, as in clinical practice, "what works" may indeed be the appropriate validation criterion.

Do we expect that research will provide simple or unambiguous answers to questions concerning the role of social learning in sustainable development? Of course not. The most relevant empirical studies that now exist (for example, Eder's [1987] work on the evolution of democracy, Hall's [1989] account of the spread of Keynesianism, Cooper's [1989] history of the development of public health practices, Peter Haas's [1990] work on the Mediterranean Action Plan, Adler's [1992] examination of the evolution of the idea of nuclear arms control, and Kai Lee's [1993] path-breaking work on the Columbia Basin) all point toward a fundamentally messy, contingent, and ambiguous intermingling of knowledge, power, interests, and chance in the workings of the world.[28] These studies nonetheless provide what we have found to be useful beginnings for efforts to understand and manipulate long-term improvements in the management of sustainable development policies. But to the extent that such improvements depend on the broad spread of images, norms, knowledge, and behavior through society at large, other models are also required. An evolutionary perspective that focuses on the population of things learned may be most suitable, but existing work in this area scarcely goes beyond provocative analogy. The immediate need is to increase the stock of empirically rich case studies informed but not constrained by today's body of theory relevant to social learning. The essays in this volume provide a welcome contribution to the task at hand.[29]

Acknowledgments

This chapter is a revision of contribution number I.6 (ver 2) to the project on Social Learning in the Management of Global Environmental Risks. It was partially supported by the National Science Foundation (through Grant number SES-9011503), the John D. and Catherine T. MacArthur Foundation, and the Social Science Research Council through its Committee for Research on Global Environmental Change. Members of this committee provided helpful feedback on an earlier version of the paper. For valuable discussions, criticism, and encouragement, we are indebted to David Ballon, Peter Cebon, Angela Liberatore, Vicki Norberg-Bohm, Douglass North, Barbara Rogoff, Thomas Schelling, and Nina Tannenwald.

NOTES

1. Our efforts thus complement those of Kai Lee (1993), who, in his extended case study of social learning in the sustainable development of the Columbia Basin, touches on many of the same themes we address here. Lee's work explores one particular view of how science and politics interact in the creation of policy. Although we are sympathetic with his approach, we attempt here to sketch a somewhat broader theoretical perspective within which his approach stands out.

2. Concise discussions of this model can be found in Elster (1986), Harsanyi (1986), and Allison (1971). Axelrod (1984), Taylor (1987), Rapoport and Chammah (1965), and Luce and Raiffa (1957) consider interactive decision problems from a rational-actor perspective.

3. Gardner (1987) provides a brief history of the behaviorist movement. Its furthest imperial expansion was Skinner (1957); the first rationalist counterattack was Chomsky (1959). Skinner (1971) is a clear popular treatment.

4. We will denote this body of theory as Social Learning Theory (capital letters), to distinguish it from the general topic of social learning.

5. Scripts can be of two kinds: episodic (a sequence of events described as a single experience [e.g., "Chamberlain at Munich"]), and categorical (generalization from common features [e.g., "appeasement only encourages aggressors to make more extreme demands"]) Abelson (1976).

6. Douglas (1986) argues that functional explanation requires that the benefits be unintended, indeed unperceived, by the individuals participating in the social form. Elster (1989) presents a skeptical view, arguing that even if norms are important determinants of individual behavior, they need not be "supraindividual entities."

7. There is, though, a recent body of empirical research on norms, based mostly on experimental studies in small groups, that focuses on norm creation and evolution and that is consequently not functionalist in orientation (e.g., Bettenhausen and Murnighan 1985; Feldman 1984; Handel 1979; Opp 1982).

8. Cole and Scribner (1974) provide a useful historical review of cross-cultural research on cultural determinants of cognition.

9. Hedberg (1981, p. 6) states that organizational learning studies mostly avoid the question of relationships between group and individual learning, using organization-level theoretical constructs but being based on observations of individuals.

10. Kimberley (1981); Rogers and Shoemaker (1971); Heclo (1974) use a contagion model to explain innovations in social policy. An early literature treating the spread of rumors with epidemic models was summarized by Dietz (1967).

11. Friedmann (1987) traces this "social engineering" view from Saint-Simon and Comte to modern muted strains in Simon and Dunn.

12. For example, Wynne's (1976) study of Barka's J-radiation demonstrating a consensus made and kept by highly selective use of evidence; Collins and Pinch's (1978) study of parapsychology demonstrating the force of unquestionable assumptions; and Frankel's (1976) demonstration that individuals' acceptance of particular claims can depend on their social position (e.g., outsiders take more risks).

13. This view, called *decisionism*, is from Shklar (1964).

14. Hall's presentation implies that he is distinguishing progressiveness in policy from policy that simply responds to present interest alignments, rather than asserting that policy learning is sluggish.

15. This approach has strong parallels to Haas's "epistemic community" work on environmental policy.

16. For example, Nye (1987); E. Haas (1990); Bennett (1990). This view fits neatly into the "incomplete learning cycle" of March and Olson (1976).

17.This is the approach of sociobiology, though, which asserts that some human behavior is genetic in origin. Wilson (1978).

18. Astley (1985), Carroll (1984), and Singh and Lumsden (1990) provide surveys of organizational ecology, distinguishing three levels of analysis: the organization, the population (of similar organizational forms), and the community.

19. Following Dawkins's usage (1976) for "units subject to imitation."

20. Emanuel Adler has been among the most careful and energetic proponents of this connection. See, for example, Adler and Crawford (1991) and Adler (1992).

21. These forms of propagation have been studied with various contagion models drawn from epidemiology (e.g., Cavalli-Sforza and Feldman [1981], and the studies reviewed by Levitt and March [1988]).

22. This applies not just to clearly evaluative phenomena such as attitudes, but also to names, definitions, and facts. For example, the concepts *racism* or *sustainable development,* or the fact that the "United States, with 5% of the world's people, consumes 25% of its energy" are learnable things in which cognition, evaluation, and affect are tightly bundled.

23. This argument is an inversion of Fleck's notion of a "thought community," with a core of initiates and a soft periphery of adherents who take the ideas literally and unquestioningly. In Fleck's view (1979), movement originates from the center and ossification occurs on the rim, but the reverse is also possible.

24. This has been most clearly recognized in the descriptions of science of Lakatos (1970) and Laudan (1977).

25. Friedmann (1987), for example, argued that significant social learning only occurs in small, task-oriented groups whose dynamics are not reducible to individuals' characteristics; that the learning is embodied in group relationships that are lost when the group dissolves; that learning occurs primarily through dialogue; and that such learning groups discover their objectives in the course of action.

26. Bateson (1942b) calls these classes proto and deutero-learning. Douglas (1986, p. 55), in a related comment, asserts, "Similarity is an institution," at least partly socially constructed.

27. Braybrooke (1987) provides a clear delineation of the naturalistic, interpretative, and critical approaches to social theory.

28. See also Kitschelt's (1986) policy history of the fast breeder reactor.

29. We are engaged in what we hope will be a complementary study focused on social response to global environmental risks.