The Challenge of Learning for Sustainability: A Prolegomenon to Theory

Adam Douglas Henry[1]
Sustainability Science Program
Harvard University
Cambridge, MA

Abstract

A core question of sustainability science asks how and why human agents learn to deal effectively with complex problems. "Learning" refers to the process by which actors assimilate information and update their cognitions and behavior accordingly. Successful learning plays a vital role in our ability to achieve sustainability, and yet this process is poorly understood. Commonly-employed perspectives on learning tend to differentiate along two dimensions: the mechanism of learning (social versus individual learning) and the properties of the information being learned (empirical versus normative knowledge). This yields four ideal types of learning that correspond to a central challenge of learning for sustainability. An integrated framework that transcends all of these perspectives is needed. Such a framework is proposed here, and includes four essential features: the structure of internal belief systems, the role of social networks in shaping knowledge, the role of knowledge in shaping networks, and the role of individual experience in the learning process. This framework is introduced as a prolegomenon (a preface to more detailed and exhaustive theoretical development) to facilitate the development of better theories and empirically-testable models of learning for sustainability.

Keywords: *sustainability, social learning, individual learning, environmental policy, social networks, cognition*

Introduction

Many realms of environmental policy are characterized by political stalemate despite massive amounts of scientific evidence that policy change is required. This is often framed as a problem of learning (Sabatier and Jenkins-Smith 1993, 1999; Social Learning Group 2001): How can decision-makers assimilate emerging and uncertain information about complex issues, and come to an agreement on appropriate strategies to resolve salient problems? The emerging field of sustainability science provides a fresh perspective on learning because of its focus on several major learning challenges in environmental policy and sustainable development (Kates et al. 2001; Clark 2007). Scholars in this community generally agree that learning is a critical hinge for sustainability (Parson and Clark 1995), but how we get there is another problem. So far there has been no systematic treatment of learning for sustainability. Despite some attempts to outline a comprehensive research program (e.g., Parson and Clark 1995; Bennett and Howlett 1992), the development of strategies to promote learning for sustainability remains an elusive goal.

Many scholars recognize the need for institutions that promote learning in the face of complex and uncertain problems. A growing literature on "collaborative policy," for example, argues that networks spanning otherwise fragmented groups of stakeholders promote an effective exchange of information and the learning of common worldviews (Schneider et al. 2003). However, there is sparse evidence that collaborative institutions and the social networks they produce actually promote learning and improved outcomes (Lubell 2004). This underscores a central problem with the literature on institutional design to promote sustainability. The process of learning is often treated as a black box, and the design of strategies to promote learning are thus based primarily on anecdotal evidence rather than lessons from theoretically-grounded and empirically-based models. A better understanding of how and why agents learn, including a detailed map of the parameters that influence this process, is a prerequisite for thinking about the types of institutions that are needed to promote learning. Understanding which types of actors are most likely to engage in constructive discussion, for example, will inform decisions regarding whom to invite to participate in a shared learning space, such as a scientific assessment process or a collaborative planning effort. These decisions must be based on a stronger theoretical and empirical understanding of how and why learning occurs. In order to accomplish this, however, we need an integrative framework that resolves the confusions and contradictions that often surround the study of learning.

The purpose of this paper is to make an initial foray into the development of such a framework. A fundamental problem, and the first obstacle to be addressed, is that there is no commonly accepted definition of learning within the sustain-

ability science community. This is largely a consequence of the interdisciplinary nature of the field. A diverse set of sister disciplines inform practical problems of sustainable development, but these disciplines tend to focus on complementary but distinct problems of learning. In economics and political science, for example, learning is often raised in the context of strategic interaction—agents are faced with the problem of "learning" optimal strategies based on experience and their beliefs about how others will behave (Camerer 2003; Axelrod 1984). The psychology literature focuses more on the internal mental processes of individual agents, and has made major advances in understanding how various types of innate cognitive biases influence the learning process (Tversky and Kahneman 1974; Kahneman 2003). This has generated an impressive body of work that has begun to be applied to explain fundamental questions of social behavior in a field known as social cognition (Kunda 1999). Other scholars view learning as an evolutionary process that involves the adoption (or "learning") of behavior and cultural traits across multiple generations (Boyd and Richerson 1985; Richerson et al. 2002; Richerson and Boyd 2005). The topic of social learning has also been addressed by philosophers such as Habermas (1996), and has gained attention in the literature on program evaluation since Campbell's (1969) work on "reforms as experiments" (Dunn 1998).

All of these perspectives are relevant to learning in the context of environmental policy and sustainable development. Their relative importance, however, depends on the particular problem under consideration. The sustainability science literature raises at least four key challenges of learning for sustainability, each of which suggests an idealized definition of learning and a unique set of variables to be considered. These challenges are discussed below. In order to develop a common language, it is necessary to transcend these individual perspectives on learning and create a synthesis that will lead to synergies across several vibrant intellectual communities that rarely build upon one another.

In addition, it should be noted that conflicts of interests often play a key role in driving political conflict over issues of sustainability. Policy choices commonly produce actual winners and losers, and although some of the conflict resolution literature focuses on social learning (Daniels and Walker 2001), there are scholars who argue that social learning cannot help to resolve these conflicts of interests (Schnaiberg and Gould 1994). Nonetheless, a learning perspective is still useful for those who think primarily in terms of interests. Understanding such conflicts, for example, depends in part on understanding how competing parties learn about the incentives and potential strategies of their opponents. Learning about the underlying details and scope of interests may also facilitate conflict resolution, since an improved understanding of opposing policy positions can serve as a basis for compromise or mutually-beneficial policy choices. Finally, interests most likely have a causal relationship with learned cognitions and behavior, such as values, beliefs, policy preferences, and norms. Although in some contexts interests may be difficult or impossible to change through a learning process, they do often bear a close relationship with valid objects of learning. Thus, a "conflict of interests" perspective can be enhanced by thinking about how learning influences the stability of interests and resultant political conflicts.

The next section addresses the problem of developing a common definition of learning. I then move to a discussion of the building blocks of an integrated framework, emphasizing in particular the role of social networks and the structure of agents' internal belief and value systems. I conclude that models of learning must explicitly address at least four key questions based on the particular learning phenomenon to be explained. These questions collectively address the co-evolutionary relationship between networking behavior and knowledge, the role of internal cognitions and cognitive bias, and the influence of individual experience in the learning process.

Defining the Challenge

Given the multiplicity of perspectives on learning, any serious treatment of this phenomenon must begin with a clear definition of terms. One of the most important tasks is to identify the appropriate unit of analysis—should we study learning at the level of societies (as in the Social Learning Group 2001), organizations (Levitt and March 1988), coalitions of policy advocates (Sabatier and Jenkins-Smith 1999), or individuals? These are some examples of units of aggregation at which learning occurs, although the sharp differences in how scholars answer the question "who learns" remains a major source of confusion in the literature (Bennett and Howlett 1992; Parson and Clark 1995).

This paper emphasizes learning by individual agents. An agent-based approach allows for some flexibility in language, since agents could represent any unit of aggregation ranging from people to nation-states. The best starting point, however, is to think of agents as individual people. This is because models of learning by individuals may be aggregated to higher levels of analysis, but the converse is not necessarily true. For example, a model of individual learning can reveal much about societal learning but studies of societal learning hide much of the richness of learning by individuals.[2] It is therefore preferable to view learning agents as individual people, although the general methods discussed here are applicable to a wide variety of conceptions of "who learns."

What is Learning?

"Learning" is a word that comes in many flavors. The most common usage refers to the process by which individuals come to understand the realities of a complex and uncertain world. Thus, economists may speak of individuals experimenting with strategies or policies and "learning" those that yield the most desired outcomes. Although this is a useful definition of learning, it is incomplete because it provides no guidance for understanding learning under conditions where there is a fuzzy or unknown reference point for reality. School children, for example, learn both facts (when was the First World War fought) as well as cultural interpretations of facts (why was the First World War fought). These interpretations of facts can be biased, impossible to falsify, or simply incorrect. But they can be "learned" in the sense that they are assimilated as truthful knowledge and impact consequent behavior. Related to this, consider also the transmission of ideas or knowledge with no empirical component whatsoever. Values can be powerful drivers of human behavior and must similarly be learned, whether at church, around the dinner table, or on the street.

How does learning relate to sustainability? Problems of sustainability are generally defined as those that involve conflicts between enhancing the well-being of humans, protecting the integrity of ecological systems, and balancing these often-conflicting goals in the long term (Kates et al. 2001; Schellnhuber et al. 2004; Clark 2007). Learning takes on many complementary meanings in the context of sustainability (see Andersson et al. 2008 for examples), but the fundamental importance of learning derives from two key problems. First, sustainability means different things to different people. Thus, engaging in sustainable development requires the learning of shared goals and appropriate strategies for achieving these goals. Second, given a shared meaning of sustainability, it is not always clear how to get there. Problems of sustainability typically arise from the complex interactions between social and ecological systems (Liu et al. 2007), and we often lack the information necessary to understand how these systems will react to any particular policy choice. This information must somehow be learned.

In this realm, learning as the accumulation of truthful knowledge about the world is an overly constraining and narrow definition, and leads to a smaller class of models than is needed to address problems of sustainability. Given a particular learning problem, it is more important to clearly define the type of cognition or behavior of interest—for example, how much an individual is willing to pay to protect endangered wildlife habitat (a value), how much an individual actually contributes to the purpose of wildlife conservation (a behavior), or a person's conviction that the money they spend on habitat protection will actually result in the desired effect of protecting biodiversity (a belief). These may be considered individually or in tandem.

Given one or more cognitions or behaviors of interest, learning is any change in these cognitions/behaviors as a result of new information or experience on the part of the focal agent. This definition does not measure learning in relation to some objective truth to be discovered. Thus, learning can be perverse in the sense that it leads an agent away from the "right" answer or optimal strategy.[3]

This definition is broad by design. When we begin to consider learning in the context of sustainability, it suddenly becomes important to consider a wide range of learned behaviors and cognitions (some with and some without clear reference points for "truth") that have a major impact on sustainability. These are generally articulated as four specific challenges; while some of the research communities that have evolved around sustainability explicitly discuss the challenge in terms of learning, others do not. I argue that the various perspectives on learning generally differentiate along two lines: the type of thing being learned, and the mechanism of learning. I turn first to a brief overview of the relevant learning challenges.

Challenge 1: Understanding Complexity

Many environmental problems are extremely complex. Oftentimes, even when there is a commonly accepted policy goal in mind (e.g., prevent severe anthropogenic interference in the climate system), the complexity and uncertainty embedded within the systems we are dealing with prevent us from knowing exactly how to accomplish stated goals. Thus, a major learning challenge is to understand the systems we wish to influence through policy choices—this allows for the intelligent matching of policy options and desired outcomes. This sense of learning presumes that there is some correct answer that may be uncovered by clever experimentation and systematic scientific inquiry. Strategies to accomplish this include policy experimentation (Campbell 1969), adaptive management and "learning by doing" (Holling 1978), and the production and sharing of knowledge through a scientific assessment process (Mitchell et al. 2006).

Challenge 2: Attenuating Normative Belief and Value Conflict

In addition to scientific complexity, environmental issues are often characterized by deeply-rooted ideological conflict. This implies the importance of another sort of learning. Competing values and normative beliefs frequently play a prominent role in driving debates over environmental policy (Lackey 2006; Sabatier and Jenkins-Smith 1993, 1999), indicating a need for decision-making processes to foster learning across competing groups of stakeholders. In this

sense, the objects of learning often include ideologies, values, or other types of cognitions with a heavy normative component. Learning is thus the process by which actors come to consensus on shared problems to be addressed and the set of acceptable strategies to be employed. This sort of learning is important not only to ensure more "enlightened" policy choices for sustainability, but also to prevent the emergence of clear ideological losers. This is critical since ideological losers are apt to stalemate the system using inside political strategies such as litigation and venue-shopping, or outside political strategies such as protests and media.

Challenge 3: Linking Knowledge with Action

Another key challenge of learning is the question of how to move knowledge from those who produce relevant information (agents of knowledge) to those who translate this knowledge into actual policy choices (agents of action). Although making optimal policy choices in the face of complexity depends on understanding that complexity, the agents who are on the front lines of meeting this challenge are not the agents responsible for making commensurate policy decisions.

The general problem of linking knowledge with action is the subject of a rich literature in its own right (see van Kerkhoff and Lebell 2006 for a detailed review). The general problem is related to the second challenge, although in this case we are concerned with the diffusion of empirical beliefs across communities that traditionally have major frictions in communication or lack opportunities to learn from one another. Much of this literature focuses on strategies to match the "supply" of scientific information with the "demands" of decision-makers (see McNie 2007), and understanding how information can be more effectively transmitted between these communities (Burns and Dietz 1992). One key finding is that information must be perceived as salient, timely, and legitimate (Cash et al. 2003; Mitchell et al. 2006).

Challenge 4: Producing New Values for Sustainability

Perhaps the most daunting challenge comes from the realization by many scholars that the values needed to support a sustainability transition are not prevalent within contemporary society (Leiserowitz et al. 2006; Dietz et al. 2005). These values must somehow be learned. This process is similar to "producing" knowledge, but in this case there is no clear reference point for the true value. Otherwise, this is conceptually similar to the first challenge—that we are in need of finding and producing values and normative beliefs that are compatible with sustainability.

An Idealized Typology of Learning

The four challenges discussed above have a strong linkage to learning processes since they all deal, in some way, with how to promote changes in environmentally-relevant cognition and behavior. The problem is that these challenges tend to focus on distinct aspects of the overall learning process. Generally speaking, these perspectives differentiate along two dimensions: the mechanism of learning and the type of knowledge that is learned. This yields four idealized types of learning, which are summarized in Figure 1.

		MECHANISM OF LEARNING *Where does knowledge originate?*	
		Individual learning	**Social learning**
SUBJECT OF LEARNING *What characterizes learned knowledge?*	**Empirical**	*Learning by doing; interpreting feedback from the environment to discover an optimal choice.* **Major perspectives:** Learning in game theory; policy experimentation and adaptive management; single-loop learning. **Associated challenges:** What is true today may not be true tomorrow; individuals have limited cognitive ability to interpret input and deal with complexity.	*Diffusion of empirical knowledge or empirically-observed behaviors amongst multiple actors.* **Major perspectives:** Linking knowledge with action; the two communities perspective. **Associated challenges:** Friction exists between scientists (agents of knowledge) and decision-makers (agents of action). It is not clear how to effectively link knowledge with action, and how to link action with knowledge.
	Normative	*Production of new values and worldviews.* **Major perspectives:** Deliberation theory; double-loop learning. **Associated challenges:** Existing values are incompatible with the needs of a sustainability transition.	*Value and worldview diffusion across agents, over space and time.* **Major perspectives:** The advocacy coalition framework; cultural evolution. **Associated challenges:** Normative beliefs and values are important drivers of environmental conflicts.

Figure 1. Four Idealized Types of Learning

The difference between individual and social learning is perhaps one of the most fundamental dimensions along which perspectives on learning are differentiated in the literature. *Individual learning* captures situations where agents are primarily concerned with learning through experience and interpreting payoffs from their environment. Individual learning processes correspond best with situations that involve the production of knowledge through experimentation or experience—that is, with Challenges 1 and 4. Although the production of new knowledge is rarely done individually, the assumption here is that knowledge is produced through experience and inquiry, rather than taught through lines of social interaction.

Social learning is the process by which agents adopt cognitions and behaviors from their social environment. In social learning situations, the relevant knowledge and behaviors already exist; the question becomes how these objects of learning are diffused though a social network that provides opportunities for mutual observation, imitation, information sharing, and learning. This is often referred to as social influence, and has been applied empirically in situations as diverse as how primates learn to use tools for food gathering (Whiten 2000) to how political beliefs are diffused through networks of friends (Johnson and Huckfeldt 2005). One of the fundamental differences between social and individual learning is that individual learning is about knowledge production, whereas social learning is about copying knowledge through imitation or persuasion.

Another distinguishing feature of these challenges is the type of knowledge that is learned. Challenges 1 and 3 focus on the learning of *empirical knowledge*, where agents are primarily concerned with reducing uncertainty and increasing the overall level of knowledge about the parameters of a specific problem. Empirical knowledge refers to facts or realities that may ideally be discovered through experimentation or direct observation. They include, for example, instrumental beliefs regarding how to achieve a specific goal or avoid unwanted consequences, such as the belief that touching a hot stove leads to a burned hand or that the emission of greenhouse gases can produce changes in the global climate. Empirical knowledge also includes information about the norms and beliefs of others in a purely descriptive way, which helps agents to better predict the likely behaviors and outcomes of their own strategies or choices. A key feature of most empirical learning models is the existence of an underlying payoff function that signals agents when they are closer to or further from the "correct" answer. This payoff function reflects the existence of an optimal belief or behavior, even if this optimum is difficult to discover due to noisy signals, inadequate information, or multiple local optimum.

Normative knowledge is the primary object of learning in Challenges 2 and 4, where there is no clear, intersubjectively reliable reference point for the correct set of beliefs or behaviors. Normative learning includes changes in values or behavioral norms as a result of empirical observations of the surrounding world. This type of knowledge tends to be highly stable and deeply embedded, and it may be that this learning occurs primarily in youth or when an agent is exposed to an extreme or horrifying event.[4] Normative knowledge must be learned nonetheless, and insofar as normative cognitions are relevant to sustainable development it is important to include this sort of knowledge in theories and models of learning. From a modeling standpoint, one of the key differences between normative and empirical learning is the role of an underling payoff function in guiding belief revision. When the object of learning is normative beliefs, this payoff function is either non-existent, or it is poorly-behaved in the sense that agents have highly dissimilar payoff structures that reward changes in their normative beliefs. Since most agents are likely to have strong accuracy goals (i.e., they seek a correct answer that is signaled by higher payoffs), these payoff structures may lead to increased heterogeneity in terms of normative beliefs and values.

The Need for a Combined Perspective

None of the typologies discussed above are entirely adequate. First, drawing a strict dichotomy between social and individual learning ignores the important insight that human behavior is both individually determined and socially embedded (Granovetter 1985). For example, models of "pure" social learning that focus solely on how knowledge diffuses in networks of social interaction tend to view agents as mere social automatons who blindly imitate the cognitions and behaviors of those around them. Models of pure individual influence assume that learning occurs within a social vacuum, and tend to idealize the ability of individual agents to act independently of their social environment. Especially in situations involving extremely complex information, individual learning models tend to over-emphasize the ability of agents to interpret raw information on their own.

The bottom line is that an agent's social network is a critical element of behavioral and cognitive change, but agents also rely heavily on accumulated individual experiences and direct observation. Thus, models of learning that simultaneously account for the role of social networks and experience can avoid conceptualizations of learning that are either over-socialized or under-socialized (Parson and Clark 1995). There are some models that take a step in this direction, such as Camerer's (2003) model of experience-weighted attraction.

Second, making the distinction between empirical and

normative objects of learning also is an artificial dichotomy. Much of our perception is theory-laden, thus we cannot properly understand the production of scientific knowledge without considering the role that normative theories play in our interpretation of real-world events (Gould 1986). Considering how empirical knowledge is learned independently of normative ideas requires a strict positivist approach that is unrealistic for most applications within issues of environmental policy (Rosa 1998). Making policy for sustainability, for example, requires an answer to the question, "What is to be sustained?" The answer often implies a direct trade-off between the things to be sustained (such as environmental protection and economic growth) and implies that normative values are an inherent component of policy choice.

In order to properly understand learning processes and move towards addressing these challenges, we are in need of a general framework of learning that transcends the individual perspectives discussed above. But how do we get there? The different perspectives of learning outlined here provide the building blocks of a general framework, discussed in the next section.

Towards an Integrated Framework

There are at least four building blocks to a general model of learning that transcends the idealized views discussed above. These building blocks, depicted in Figure 2, are the various processes that tie into explaining how a particular cognition or behavior is ultimately learned. The building blocks ultimately revolve around the cognitions and behavior of an individual agent—in particular, those changing cognitions and/or behaviors that we seek to explain as a result of learning.

These building blocks include the internal cognitive structure of agents, their outward networking behavior, the role of social influence, and the role of individual experience. These comprise a general framework that may be refined into more specific models depending on the context in which learning takes place. Some learning problems, for example, may necessitate complex models of cognition and behavior whereas others require relatively simple models. These building blocks are summarized in the following sections.

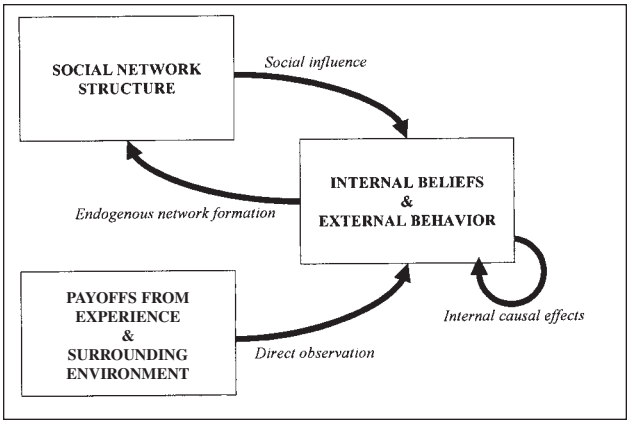

Figure 2. A Basic Framework of the Learning Process

How are Cognitions Structured Internally?

In studying how and why agents learn, it is common to focus only on external factors such as the role of social influence and direct observation. But agents' internal cognitions also have a causal structure, and thus it is important to consider the role of additional "internal" factors that influence the learning process. There may be a disconnect between the cognition or behavior that is ultimately changed as a result of the learning process, and the cognitions or behaviors that are influenced by the proximate cause of learning—namely, one's social network or payoffs from their beliefs or behaviors. A key part of learning is to understand the causal links between the objects of learning and other intermediate cognitive and behavioral factors.

Perhaps the simplest example of this is the problem of learned behavior. Consider the example of why individuals are beginning to purchase more fuel efficient vehicles (a behavior). There are several general possibilities. The first possibility is that this behavior is learned through simple imitation. A focal agent buys a hybrid, for example, because both of her neighbors own hybrid vehicles. A second possibility is that this behavior is learned indirectly through belief change. That is, the fact that your neighbors both own fuel efficient vehicles exposes you to new information about how much money you might save on gasoline (Denrell 2008). It is this new realization (a belief) that is learned, and the best way to act on this learned belief is to purchase a more fuel efficient vehicle—this behavior is independent from what your neighbors do. This possibility implies the need to consider a more complex structure of cognition and behavior. In the case of imitation it is sufficient to consider the relevant behavior only; in the latter example, however, it is necessary to consider beliefs, behaviors, and a causal link that relates the two together. Which of these models is correct has profound implications for efforts to promote more sustainable behavior.

This example illustrates the need to be clear about the set of cognitions and behaviors that are learned, or that somehow influence the thing to be learned. The simple possibilities outlined above suggest at least two relevant phenomena: a belief about the money that can be saved from purchasing a hybrid car, and the behavior that comes from acting upon this belief. Other possibilities would include a more diverse set

of cognitions, such as preferences (beliefs about saving money causes fuel-efficiency to be activated as a preference, which activates buying behavior) or values (environmentalism spreads across neighbors, which is reflected in learned behavior). A model of learning must not only outline the space of cognitions and behaviors that are relevant to the problem at hand, but it must also posit reasonable hypotheses concerning the causal structure of these cognitions.

Some problems may necessitate only simple models of internal cognitive structure. These are most likely to be problems where agents learn about very simple problems, or there is some flat, uni-dimensional characteristic to the knowledge being learned. But models of learning for sustainability should, at a minimum, address both normative and positive cognitions—in particular, empirical beliefs about causal relationships as well as underlying value structures. Frameworks such as the Advocacy Coalition Framework (Sabatier and Jenkins-Smith 1993, 1999; Sabatier and Weible 2007) or Values-Beliefs-Norms theory (Stern et al. 1999) contain hypotheses regarding the causal structure of cognitions, and their relationship to salient behaviors, that are potentially useful for this purpose. More elaborate models of cognition may also include an explanation of how interests drive behavior, or are in turn influenced by learned cognitions such as preferences and values.

Just as the appropriate model of internal cognitive structure depends on the context of the learning problem, the same model may not apply equally well to all agents within a learning space. For example, members of the general public are unlikely to learn about environmental issues in the same way as members of the policy elite. The engagement and specialization of elites within their respective issue area suggests some unique features of an elite learning model, including a more deliberate search for accurate information and a well-defined, highly integrated system of policy-relevant beliefs (Henry 2007).

What is the Role of Experience?

Some of the most sophisticated perspectives on learning have been models of individual learning, where the fundamental goal is to understand how an agent learns about some specific aspect of how the world works. The ultimate goal of this sort of learning is to understand how actions produce outcomes (or payoffs) as a function of variables describing the real world. It is therefore important to understand that world. Models of learning should address the role of individual experience and active "truth-seeking" in the learning process.

For example, most game theoretic perspectives on learning fit well within the individual learning tradition (see Camerer 2003 for examples). Although agents learn in a strategic game setting in a social environment, the learning task is for the individual agent to better understand the likely ways in which opponents will react to strategies. Thus, opponents' strategies (although they are embedded in the social environment) comprise the knowledge about which the focal agent learns.

The literatures that touch on this are immense and can be overwhelming. The obvious task, which is beyond the scope of this paper, is to find those theories of learning that are theoretically and empirically strongest, and that are most applicable to the learning problems faced in the context of sustainability.

What is the Role of Social Influence?

Learning rarely occurs within a social vacuum. This is especially true in any learning situation dealing with highly complex information that is too much for a single agent to master. Legislators, for example, often specialize in particular issue areas but also rely heavily on information from lobbyists, constituent groups, and staff. The prominent role of social influence in most learning situations underscores the need to include a clear operationalization of an agent's social environment within a model of learning. Social networks are a vital component of learning because they act as a constraint on information exchange and the pooling of knowledge at multiple scales. Multiple perspectives on a problem are useless, for example, if there is no mechanism for synthesizing these perspectives. As noted above, networks are crucial because they are a medium for information sharing, dialogue, persuasion, negotiation, and any other social process that leads to belief change or knowledge production.

Social network structure also reflects the institutions in which learning agents are embedded. In some cases, institutional arrangements directly impose structure on a network—for example, environmental policies commonly require that one or more regulatory agencies work together to implement laws. Other times, institutional arrangements impose structure indirectly by creating opportunities for linkages to form. In collective action problems, for example, rules that allow agents to openly discuss strategies and impose sanctions on one another can lead to increased levels of cooperation (Ostrom et al. 1992), thereby supporting the formation of reciprocated networks of cooperation. There are many other ways that institutional rules influence network structure, either by creating incentives or disincentives for cooperation, or mandating the formation of an ideal type of network (for example, a network that includes all relevant stakeholders in a decision process). Institutional design therefore has the potential to influence learning processes, at least in part due to the relationship between institutional rules and networks (see Ostrom 2005 for a detailed typology and discussion of institutional rules).

Social network analysis (Wasserman and Faust 1994; Scott 2000; Henry 2008) provides the conceptual tools necessary to describe an agent's opportunities for learning through the social environment, and suggests some predictions concerning the relationship between network structure, knowledge, and behavior. From an analytic standpoint, learning via social influence is essentially the same as diffusion throughout a social network. There has been a great deal of work in the networks literature focusing on how structure promotes or hinders diffusion processes (see Newman 2003 for a review). The applied settings of these models are diverse, ranging from the spread of disease to the spread of panic in a social network. Work in this vein that bears a closer relationship to learning processes investigates how network structure influences the emergence of cooperative behavior amongst a population of agents engaged in a prisoner's dilemma type of situation (Ohtsuki et al. 2006).

The core insight of this literature is that network structure strongly influences the outcomes of social interaction, including social learning. Thus, models of learning must explicitly address how the position of agents within a network influences the learning process. The detail with which these dynamics are specified depends on the needs of the model. At a minimum, however, models must specify the importance of social influence processes and (if social influence is considered to be an important component of learning) models should include a variable that captures an agent's proximate social environment.

How Do Cognitions Influence the Social Environment?

Social networks are a key component of learning, but they are rarely static over time. Despite the effect of institutions on network structure, it is unlikely that sets of institutional arrangements can ever completely predict network structure. Rather, networks are in constant flux as agents move through the spatial environment, and as new relationships are forged or old ones broken based on the predilections and biases of individual agents. There is a growing literature that seeks to identify the multiple factors driving the formation and evolution of social networks (Snijders 2005). Some of the common results are that actors deliberately embed themselves within special types of network configurations, such as closed triads or reciprocated relationships (Robins et al. 2007).

These factors can potentially be viewed as exogenous constraints on learning, and therefore fit into the above set of variables that describe network structure. However, an agent's cognitions and behavior (the potential objects of learning) also influence how she positions herself within networks. Thus, networks are also products of learning due to the inherent co-evolution of individual attributes, such as beliefs and values, and larger network structures (Lazer 2001). While networks act as a constraint on what one agent *can* learn, what agents *have* learned influences how they structure their social environments. A better understanding of learning processes depends on considering how knowledge and networks interact and change over time.

Models of learning must therefore address one specific aspect of network change: how the objects of learning (cognitions and behaviors) lead agents to restructure their proximate social networks. For example, there is a growing body of empirical evidence that networks are structured by belief homophily, or the tendency for agents to seek out network partners that share their underlying belief systems (Weible and Sabatier 2005; Weible 2005; McPherson et al. 2001). If agents have the freedom to create or reject relationships on the basis of belief homophily, then networks will tend to fragment along ideological lines. A simple model of how this occurs (or a convincing explanation of why it will not occur) is essential to understanding the constraints that agents face in learning from their social environments.

Conclusion

The preceding discussion sketches a general framework to guide inquiry into questions of learning for sustainability, but moving towards a coherent theory of learning requires formulating models that are precise and empirically testable. This paper is meant to be a prolegomenon to a theory of learning—an initial foray into the subject and a starting point for more extensive theoretical development and empirical analysis. In particular, future work should begin to formulate precise models of learning that integrate the multiple categories of variables that are considered important across a variety of disconnected but complementary literatures.

This paper outlines four key inputs to these models. However, the relative importance of these inputs will depend greatly on context. Depending on the particular situation in which a model of learning is meant to be applied, social influence, internal belief systems, individual experience, and endogenous network formation may play very different roles in shaping the beliefs and behaviors of individual agents. In some contexts, simpler models of the sort that have already been developed in the literature will work. For example, a simple model of how ideas diffuse along a fixed social network is appropriate in cases where agents have a strong incentive to adopt the ideas of their network partners and little or no ability to seek information from alternative sources.

Although simpler models may be justified in certain contexts, it is important to recognize that the simpler approach is not the only possible approach. Future work will need to focus on how different contextual factors require

models that incorporate more or less detail in the four building blocks discussed above. One of the more important aspects of context will be the question of who learns. At the level of individual people, for example, different types of agents (e.g., scientists versus legislators) face very different learning problems and operate within very different learning environments. More empirical and conceptual work is needed to uncover these differences and develop an appropriate set of context-sensitive learning models.

These models should be evaluated according to their theoretical and empirical strengths, but they should also be evaluated according to the degree to which they resonate with the types of learning that are needed for sustainability. This paper outlines two major features of learning for sustainability that should be considered. In particular, models should pay attention to the empirical and normative dimensions of learned knowledge, and acknowledge both the social and individual aspects of the learning process. Ultimately, these models should help us understand how institutions may be designed to successfully meet the challenges of learning for sustainability.

Endnotes

1. E-mail: adam_henry@hks.harvard.edu
2. This is because higher units of aggregation collapse an entire social space into a single metaphorical agent. If agents are firms, for example, then the complex social interactions and associated learning within the firm becomes a part of agents' internal cognitive processes.
3. The possibility of perverse learning is often discounted under the assumption that agents are rational actors focused on optimizing some set of payoffs, or at least use learning heuristics that lead to better outcomes (Dietz and Stern 1995).
4. In addition, Habermas (1971, 1991, 1993) argues that deliberative processes may be a mechanism by which normative learning can occur later in life or in the absence of extreme events.

Acknowledgements

This research was supported by a Giorgio Ruffolo Doctoral Fellowship in Sustainability Science, the Center for International Development at Harvard University, and by a dissertation grant through the Sustainable Transportation Center at the University of California, Davis. I also wish to thank the anonymous reviewers for their insightful comments on an earlier draft of this paper.

References

Andersson, K., M. Burns, M. Bursztyn, A.D. Henry, A. Laudati, K. Matus, and E. McNie (eds.). 2008. *The Ruffolo Curriculum on Sustainability Science: 2008 Edition*. CID Graduate Student and Research Fellow Working Paper No. 32, Center for International Development at Harvard University, December 2008.

Axelrod, R. 1984. *The Evolution of Cooperation*. New York: Basic Books.

Bennett, C.J. and M. Howlett. 1992. The lessons of learning: Reconciling theories of policy learning and policy change. *Policy Sciences* 25, 275-294.

Boyd, R. and P.J. Richerson. 1985. *Culture and the Evolutionary Process*. Chicago, IL: University of Chicago Press.

Burns, T.R. and T. Dietz. 1992. Technology, sociotechnical systems, technological development: An evolutionary perspective. In M. Dierkes and U. Hoffmann (eds.), *New Technology at the Outset: Social Forces in the Shaping of Technological Innovation*, 206-238. Frankfurt am Main: Campus Verlag.

Camerer, C. 2003. *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton University Press.

Campbell, D.T. 1969. Reforms as experiments. *American Psychologist* 24, 4, 409-429.

Cash, D.W., W.C. Clark, F. Alcock, N.M. Dickson, N. Eckley, D.H. Guston, J. Jäger, and R.B. Mitchell. 2003. Knowledge systems for sustainable development. *Proceedings of the National Academy of Sciences* 100, 14, 8086-8091.

Clark, W.C. 2007. Sustainability science: A room of its own. *Proceedings of the National Academy of Sciences* 104, 6, 1737-1738.

Daniels, S.E. and G.B. Walker. 2001. *Working Through Environmental Conflict: The Collaborative Learning Approach*. Westport, CT: Praeger.

Denrell, J. 2008. Indirect social influence. *Science* 321, 47-48.

Dietz, T. and P.C. Stern. 1995. Toward realistic models of individual choice. *Journal of Socio-Economics* 24, 261-279.

Dietz, T., A. Fitzgerald, and R. Shwom. 2005. Environmental values. *Annual Review of Environment and Resources* 30, 335-372.

Dunn, W.N. (ed.). 1998. *The Experimenting Society: Essays in Honor of Donald T. Campbell. Policy Studies Review Annual*, Volume 11.

Gould, S.J. 1986. *The Mismeasure of Man*. New York: W.W. Norton.

Granovetter, M. 1985. Economic action and social structure: The problem of embeddedness. *American Journal of Sociology* 91, 3, 481-510.

Habermas, J. 1971. *Knowledge and Human Interests*. Boston, MA: Beacon Press.

Habermas, J. 1991. *Moral Consciousness and Communicative Action*. Boston, MA: Beacon Press.

Habermas, J. 1993. *Justification and Application: Remarks on Discourse Ethics*. Cambridge, Massachusetts: The M.I.T. Press.

Habermas, J. 1996. *Between Facts and Norms: Contributions to a Discourse Theory of Law and Democracy*. Translated by W. Rehg. Cambridge, MA: MIT Press.

Henry, A.D. 2007. Simulating the evolution of policy-relevant beliefs: Can rational learning lead to advocacy coalitions? In E. Ostrom and A. Schlüter (eds.), *The Challenge of Self-Governance in Complex, Globalizing Economies*, 135-160. University of Freiburg Institute of Forestry Economics Working Paper 47e.

Henry, A.D. 2008. Social network analysis: An emergent method. In L. Kalof, A. Dan and T. Dietz, *Essentials of Social Research*, 138-140. Berkshire, UK: Open University Press/McGraw Hill Publishers.

Holling, C.S. 1978. *Adaptive Environmental Assessment and Management*. New York: John Wiley & Sons.

Johnson, P.E. and R. Huckfeldt. 2005. Agent-based explanations for the survival of disagreement in social networks. In A.S. Zuckerman (ed.), *The Social Logic of Politics*, 251-268. Philadelphia: Temple University Press.

Kates, R.W., W.C. Clark, R. Corell, J.M. Hall, C.C. Jaeger, I. Lowe, J.J. McCarthy, H.J. Schellnhuber, B. Bolin, N.M. Dickson, S. Faucheux, G.C. Gallopin, A. Grübler, B. Huntley, J. Jäger, N.S. Jodha, R.E. Kasperson, A. Mabogunje, P. Matson, H. Mooney, B. Moore III, T. O'Riordan, and U. Svedlin. 2001. Sustainability science. *Science* 292, 641-642.

Kahneman, D. 2003. A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist* 58, 9, 697-720.

Kunda, Z. 1999. *Social Cognition: Making Sense of People*. Cambridge, MA: The MIT Press.

Lackey, R.T. 2006. Axioms of ecological policy. *Fisheries* 31, 6, 286-290.

Lazer, D. 2001. The co-evolution of individual and network. *Journal of Mathematical Sociology* 25, 1, 69-108.

Leiserowitz, A.A., R.W. Kates, and T.M. Parris. 2006. Sustainability values, attitudes, and behaviors: A review of multinational and global trends. *Annual Review of Environment and Resources* 31, 413-444.

Levitt, B. and J.G. March. 1988. Organizational learning. *Annual Review of Sociology* 14, 319-340.

Liu, J., T. Dietz, S.R. Carpenter, M. Alberti, C. Folke, E. Moran, A.N. Pell, P. Deadman, T. Kratz, J. Lubchenco, E. Ostrom, Z. Ouyang, W. Provencher, C.L. Redman, S.H. Schneider, W.W. Taylor. 2007. Complexity of coupled human and natural systems. *Science* 317, 5844, 1513-1516.

Lubell, M. 2004. Collaborative environmental institutions: All talk and no action? *Journal of Policy Analysis and Management* 23, 3, 549-573.

McNie, E.C. 2007. Reconciling the supply of scientific information with user demands: An analysis of the problem and review of the literature. *Environmental Science & Policy* 10, 17-38.

McPherson, M., L. Smith-Lovin, and J.M. Cook. 2001. Birds of a feather: Homophily in social networks. *Annual Review of Sociology* 27, 415-444.

Mitchell, R.B., W.C. Clark, D.W. Cash, and N.M. Dickson (eds.). 2006. *Global Environmental Assessments: Information and Influence.* MIT Press.

Newman, M.E.J. 2003. The structure and function of complex networks. *SIAM Review* 45, 2, 167-256.

Ohtsuki, H., C. Hauert, E. Lieberman, M.A. Nowak. 2006. A simple rule for the evolution of cooperation on graphs and social networks. *Nature* 441, 502-505.

Ostrom, E., J. Walker, and R. Gardner. 1992. Covenants with and without a sword: Self-governance is possible. *American Political Science Review* 86, 2, 404-417.

Ostrom, E. 2005. *Understanding Institutional Diversity*. Princeton University Press.

Parson, E.A. and W.C. Clark. 1995. Sustainable development as social learning: Theoretical perspectives and practical challenges for the design of a research program. In L.H. Gunderson, C.S. Holling, and S.S. Light (eds.), *Barriers and Bridges to the Renewal of Ecosystems and Institutions*. New York: Columbia University Press.

Richerson, P.J., R. Boyd, and B. Paciotti. 2002. An evolutionary theory of commons management. In E. Ostrom, T. Dietz, N. Dolsak, P.C. Stern, S. Stonich, and E. Weber (eds.), *The Drama of the Commons*, 403-442. Washington, D.C.: National Academy Press.

Richerson, P.J. and R. Boyd. 2005. *Not by genes alone: How culture transformed human evolution*. Chicago, IL: University of Chicago Press.

Robins, G., T. Snijders, P. Wang, M. Handcock, and P. Pattison. 2007. Recent developments in exponential random graph (p*) models for social networks. *Social Networks* 29, 192-215.

Rosa, E. 1998. Metatheoretical foundations for post-normal risk. *Journal of Risk Research* 1, 15-44.

Sabatier, P. and H. Jenkins-Smith (eds.). 1993. *Policy Change and Learning: An Advocacy Coalition Approach*. Boulder: Westview Press.

Sabatier, P. and H. Jenkins-Smith. 1999. The advocacy coalition framework: An assessment. In P. Sabatier (ed.). *Theories of the Policy Process*, 117-166. Boulder: Westview Press.

Sabatier, P.A. and C.M. Weible. 2007. The advocacy coalition framework: Innovations and clarifications. In P.A. Sabatier (ed.). *Theories of the Policy Process*. Second Edition. Boulder: Westview Press.

Schellnhuber, H.J., P.J. Crutzen, W.C. Clark, M. Claussen, and H. Held (eds.). 2004. *Earth Systems Analysis*. Cambridge, MA: MIT Press.

Schnaiberg, A. and K.A. Gould. 1994. *Environment and Society: The Enduring Conflict*. New York: St. Martin's Press.

Schneider, M., J. Scholz, M. Lubell, D. Mindruta, and M. Edwardsen. 2003. Building consensual institutions: Networks and the national estuary program. *American Journal of Political Science* 47, 1, 143-158.

Scott, J. 2000. *Social Network Analysis*, 2nd Edition. Sage Publications.

Snijders, T.A.B. 2005. Models for longitudinal network data. In P.J. Carrington, J. Scott, and S. Wasserman (eds.). *Models and Methods in Social Network Analysis*, 215-247. Cambridge University Press.

Social Learning Group. 2001. *Learning to Manage Global Environmental Risks*. Volumes 1 and 2. The MIT Press.

Stern, P.C., T. Dietz, T. Abel, G.A. Guagnano, and L. Kalof. 1999. A social psychological theory of support for social movements: The case of environmentalism. *Human Ecology Review* 6, 81-97.

Tversky, A. and D. Kahneman. 1974. Judgment under uncertainty: Heuristics and biases. *Science* 185, 1124-1131.

van Kerkhoff, L. and L. Lebel. 2006. Linking knowledge and action for sustainable development. *Annual Review of Environmental Resources* 31, 445-477.

Wasserman, S. and K. Faust. 1994. *Social Network Analysis: Methods and Applications*. Cambridge University Press.

Weible, C.M. 2005. Beliefs and perceived influence in a natural resource conflict: An advocacy coalition approach to policy networks. *Political Research Quarterly* 58, 3, 461-475.

Weible, C.M. and P.A. Sabatier. 2005. Comparing policy networks: Marine protected areas in California. *The Policy Studies Journal* 33, 2, 181-201.

Whiten, A. 2000. Primate culture and social learning. *Cognitive Science* 24, 3, 477-508.