Chapter Seven

Usable Knowledge:
Design Principles and Institutional Diagnostics

THE PURSUIT OF KNOWLEDGE about the institutional dimensions of environmental change is worthwhile as an end in itself. Yet much recent interest in this subject arises from a more applied concern with producing usable knowledge or, in other words, knowledge that will enhance our ability to control or regulate human actions that cause or threaten to cause disruptive – possibly irreversible – changes in human-dominated ecosystems (Vitousek et al. 1997), ranging from smallscale systems (e.g. local fish stocks or forests) that support limited groups of local appropriators to global systems (e.g. the Earth's climate system) that support all life on the planet (NRC 1999). We want to know how and why some local appropriators have found ways to use renewable resources in a manner that is sustainable over time while others have failed conspicuously to do so and whether we can draw lessons from these experiences that are applicable to other cases (Ostrom et al. 1999). Similarly, we want to know why some international regimes are more successful than others in solving environmental problems and what we can learn from the successful cases about the design of effective regimes to deal with a range of other problems (Young 1999a). In each case, the underlying goal is the same, the derivation of lessons from past and present experience that can help in the process of designing institutions to deal with current or future environmental changes. Accordingly, the relationship between analysis and praxis constitutes a topic of particular importance to all those seeking to improve our capacity to manage the institutional dimensions of environmental change.

Under the circumstances, the existence of a sizable gap between the world of analysis and the world of practice in this realm is cause for concern. Despite the efforts of a few individuals who have acquired the skills needed to operate comfortably in both worlds, most attempts to bridge this gap have yielded meager results. Researchers often frame their conclusions in abstract terms and publish their results in journals that practitioners seldom read. They emphasize uncertainties and are reluctant to make recommendations based on the results of their work for the handling of specific environmental problems, such as climate change or the loss of biological diversity. For their part, policymakers commonly assume that it is difficult or impossible to come to terms with specific problems in the absence of detailed and continuously updated knowledge about the state of play on individual issues. They tend to look upon researchers as armchair analysts who have little contact with the actual course of events on the ground.

What can we do to bridge this gap and, in the process, to improve the quality of the dialogue between the two communities? In this concluding chapter, I address this question first through an examination of the effort to develop design principles, which constitutes the most influential approach to the task of bringing analysis to bear on practice in this realm, and then through a discussion of institutional diagnostics, which I portray as a distinct – but by no means mutually exclusive - approach to this task. In the process, I argue that although the idea underlying design principles is a highly attractive one, there are serious limitations to what we can hope to achieve in the search for design principles applicable to a wide range of environmental concerns. The institutional diagnostics approach is less ambitious and therefore less attractive in ideal terms than the design principles approach. Even so, I conclude that the practice of institutional diagnostics is likely not only to produce results that are more usable in dealing with specific problems but also to pave the way for a productive and ongoing dialogue between researchers and policymakers in the search for institutional arrangements capable of limiting – if not eliminating - costly anthropogenic changes in human-dominated ecosystems.

1. Design Principles

The goal of developing generalizations that capture major findings about the institutional dimensions of environmental change and framing them in such a way that they can be applied to the handling of specific problems is relevant to all social settings. So far, however, those considering local common-pool resources (CPRs) have made the most sustained and productive efforts to pursue this type of reasoning. As Elinor Ostrom, the leading contributor to this stream of analysis, puts it, a design principle is "… an essential element or condition that helps to account for the success of … institutions in sustaining the CPRs and gaining the compliance of generation after generation of appropriators to the rules in use" (Ostrom 1990: 90). Basing their conclusions on a systematic

TABLE 1 DESIGN PRINCIPLES ILLUSTRATED BY LONG-ENDURING CPR INSTITUTIONS
1. Clearly defined boundaries Individuals or households who have rights to withdraw resource units form the CPR must be clearly defined, as must the boundaries of the CPR itself.
2. Congruence between appropriation and provision rules and local conditions Appropriation rules restricting time, place, technology, and/or quantity of resource units are related to local conditions and to provision rules requiring labor, material, and/or money.

3. Collective-choice arrangements Most individuals affected by the operational rules can participate in modifying the operational rules.
4. Monitoring Monitors, who actively audit CPR conditions and appropriator behavior, are accountable to the appropriators or are the appropriators.
5. Graduated sanctions Appropriators who violate operational rules are likely to be assessed graduated sanctions by other appropriators, by officials accountable to these appropriators, or by both.
6. Conflict-resolution mechanisms Appropriators and their officials have rapid access to low-cost local arenas to resolve conflicts among appropriators or between appropriators and officials.
7. Minimal recognition of rights to organize The rights of appropriators to devise their own institutions are not challenged by external governmental authorities.
Source: Ostrom 1990: 90

examination of a sizable collection of actual cases, Ostrom and her colleagues have produced a set of specific design principles illustrated by long-enduring CPR institutions (see Table 1) (Ostrom 1990, Gibson, McKean, and Ostrom 2000).

In her account of these principles, Ostrom is suitably modest about the current status of this endeavor, observing that she is "… not yet willing to argue that these design principles are necessary conditions for achieving institutional robustness in CPR settings" and that "[f]urther theoretical and empirical work is needed before a strong assertion of necessity can be made" (Ostrom 1990: 90). Nonetheless, the direction in which this line of reasoning is moving is clear.

Ostrom states plainly that she is "… willing to speculate … that after further scholarly work is completed, it will be possible to identify a set of necessary design principles" (Ostrom 1990: 90-91). To the extent that it succeeds, the implications of this effort for policymaking will be both far-reaching and easy to grasp. All those seeking to (re)form management regimes dealing with CPRs will know that the arrangements they create cannot endure or produce sustainable outcomes unless they include provisions that address the concerns identified in each of the design principles.

1.1 Distinctive Features of Design Principles

What can we say about the character and content of these design principles? For starters, it is apparent that Ostrom's design principles emanate from an analysis inspired mainly by collective-action models in contrast to social-practice models. The analysis begins with a consideration of familiar problems such as the tragedy of the commons, prisoner's dilemma, and the logic of collective action and proceeds to an analysis of "the evolution of institutions for collective action."[1] There is nothing in the generic concept of design principles that requires a commitment to collective-action models. It is perfectly possible, for instance, that others could produce a set of principles highlighting sociological conditions involving such matters as feelings of legitimacy, the role of consensual knowledge, or the sense of community required to sustain various forms of diffuse reciprocity. Yet Ostrom's emphasis on conditions relating to collective-choice procedures, monitoring, sanctions, and conflict-resolution mechanisms is perfectly understandable in terms of thinking that rests on the assumptions characteristic of collective-action models.

Still, there is an element of ambiguity about the target or goal of the exercise in Ostrom's account of design principles. Are the individual principles necessary for the attainment of "institutional robustness" in the sense of arrangements that are long enduring or capable of governing the activities of "generation after generation of appropriators" or for the achievement of sustainability evaluated in terms of the biogeophysical condition of the ecosystems in question?[2] No doubt, it is reasonable to assume that there is a close relationship between robustness and sustainability. There is little prospect that an institutional arrangement can survive over a long period of time if it fails to produce results that are sustainable in biogeophysical terms. Even so, there are good reasons to think carefully about the relationship between robustness and sustainability. If the relevant ecosystems are subject to nonlinear changes that are natural or non-anthropogenic in origin, the achievement of robustness will require a well-developed capacity to monitor and adjust to natural changes in the state of key ecosystems. If such changes can and occasionally do occur quite suddenly so that there is a real danger of serious surprises, the pursuit of robustness will require the creation and maintenance of early warning systems. And if there is substantial uncertainty about these matters, appropriators will be well-advised to devote resources to improving their knowledge base regarding the behavior of the relevant ecosystems, introducing various precautionary measures, or adopting a more complex strategy featuring both types of response.

It is worth paying explicit attention as well to the fact that the design principles approach focuses on the goal of formulating necessary conditions or identifying links that are

---

[1] . This phrase constitutes the subtitle of Ostrom 1990.

[2] . While Ostrom herself tends to emphasize the idea of long-enduring arrangements, the general literature on the tragedy of the commons focuses on the biogeophysical condition of CPRs or, more generally, on sustainability (Hardin and Baden 1977, Baden and Noonan 1998).

unconditional in nature. From the perspective of demonstrating policy relevance, this feature of the approach has obvious attractions. It means, above all, that there should be no exceptions. If the existence of active and accountable monitors or the operation of credible sanctions is a true design principle, it would be foolhardy to ignore the prescriptive implications and suppose that it is possible in specific cases to cut corners when it comes to the organization of monitoring mechanisms capable of assessing CPR conditions or compliance mechanisms capable of imposing appropriate sanctions. But several other aspects of this feature of design principles are noteworthy as well. The principles tell us nothing about sufficient conditions for the achievement of robustness or sustainability in managing human uses of CPRs. It is perfectly possible, therefore, that appropriators could adhere scrupulously to all the design principles in their efforts to construct resource regimes but nonetheless fail to achieve robust or sustainable results. The emphasis on necessary conditions suggests also that options available to those seeking to achieve robust or sustainable outcomes may be quite limited, a conclusion that will seem counterintuitive both to analysts who think in terms of equifinality and to practitioners who believe that there is more than one way to solve most problems involving human actions (Young and Osherenko 1993). Coupled with the sheer number of distinct principles, the focus on necessity may engender a certain sense of pessimism as well. Ostrom identifies seven separate principles and suggests that we might ultimately need to add a few more.[3] Under the circumstances, policymakers may conclude understandably that they are being asked to make an enormous effort to devise arrangements that fulfill a lengthy list of necessary conditions but that still offer no assurance that the results will prove to be robust or sustainable.[4]

---

[3] . To these, she adds an eighth condition relating to specifically to CPRs that are parts of larger systems (Ostrom 1990: 90).

[4] . Of course, many of the long-enduring CPR institutions that Ostrom considers in her study have evolved over time without conscious efforts on the part of their creators. In such cases, the number of

Because design principles are framed as universal propositions, they should hold across all members of the relevant universe of cases. In Ostrom's work, this universe encompasses the set of smallscale CPRs, where smallscale is defined in terms of number of appropriators as well as spatial domain and CPRs are construed as goods and services that are both non-excludable and rival or subtractable. Any situation exhibiting these characteristics in which appropriators succeed in producing results that are robust and sustainable without fulfilling one or another of the design principles would count as evidence against that principle. If this were to happen in more than the odd instance, a full-scale reassessment of the status of the relevant principle would be in order.

An obvious question, then, concerns the generalizability of these design principles to cases that lie outside the universe of smallscale CPRs. In part, this is a matter of scale (Young 1994). Can we scale up from smallscale systems to largescale systems and ultimately to global systems in thinking about design principles relating to the institutional dimensions of environmental change (see Ch. 6 for a more general account of the problem of scale)? In part, it is a matter of the nature of the problems at stake. To what extent do the members of the overall set of environmental problems - including those occurring at the local level - exhibit the defining features of common-pool resources? Understandably, Ostrom and her colleagues have argued that there are reasonable grounds for answering both these questions in the affirmative (McGinnis and Ostrom 1996, Ostrom et al. 1999). But precisely because the incentive to arrive at

---

necessary conditions may be irrelevant. But so also is any activity we might reasonably describe as institutional design.

positive answers to these questions is strong, we should examine the issues at stake in this connection with particular care.

1.2 Limits of Design Principles

The logic of design principles requires the existence of a universe of cases that is both well-defined, at least in analytic terms, and homogeneous, at least with regard to those factors relevant to the creation of regimes that are long enduring and capable of producing sustainable results. If we are unable to determine whether specific situations do or do not belong to the universe, the degree to which design principles constitute necessary conditions will be difficult to test. If, on the other hand, the boundaries of the universe are clear but the members of the universe are heterogeneous in important respects, individual design principles will be testable but they will fail to meet the test for (nearly) necessary conditions.

Is the universe of smallscale CPRs well-defined and homogenous? There is ample scope for debate regarding this question. To take a single example, it seems likely that those who think in terms of social-practice models of institutions would point to a number of factors (e.g. the presence or absence of a strong sense of community) that could affect the way individual appropriators use living resources but that are not easily captured in the ideas of non-excludability and subtractability (Singleton and Taylor 1992). But I do not wish to examine this issue in any depth here. For the sake of the argument, let us accept the proposition that smallscale CPRs are, by and large, reasonably well-defined and homogeneous. The relevant question then becomes a matter of determining the extent to which the larger class of environmental problems arising at scales ranging from the micro-level of local ecosystems to the macro-level of global ecosystems share the defining characteristics of smallscale CPRs. This is not a simple or straightforward question, a fact that should suffice to warn us against offering casual or offhand responses to it. Even so, I think it is fairly easy to show that the level of heterogeneity in the larger class of environmental problems is high enough to cast grave doubts on any assumption to the effect that what holds for smallscale CPRs will hold for other classes of problems involving environmental changes as well.

Consider several well-known analytic distinctions as vehicles for guiding our thinking about this issue. There is, to begin with, the familiar dichotomy between coordination problems and collaboration problems (Stein 1982). Whereas regimes that address collaboration problems (e.g. any situation that can be modeled as a prisoner's dilemma) leave individual participants with incentives to violate the rules or to cheat in meeting their commitments, coordination problems can be solved through the development of regimes in which none of the participants experiences any incentive to cheat or defect. In game-theoretic terms, regimes created in to deal with coordination problems are characterized by the existence of a stable equilibrium. Although it may be tempting to dismiss this finding as an interesting artifact having little relevance to the real world, the evolution of social conventions or rules that leave everyone more or less content in a wide range of situations suggests that the category of coordination problems is far from empty in empirical terms. Much the same can be said regarding situations in which there are significant conflicts of interest coupled with the existence of two or more equilibrium outcomes (e.g. battle of the sexes). Regime formation, in such situations, may be characterized by hard bargaining in which each of the parties uses a variety of tactics to persuade or compel the other(s) to settle for its preferred outcome. But once a bargain is struck, the members of the resultant regime will experience no ongoing incentives to cheat on the terms of the agreement (Schelling 1960). In all these cases, the achievement of robustness will not depend on meeting the requirements of principles dealing with monitoring and graduated sanctions. It may not even be necessary to worry about clearly defined boundaries, so long as new members are willing to accept or accede to the existing rules when they become players in an ongoing game.

Similar conclusions arise from a consideration of well-known economic constructs. Where the problem centers on persuading participants to contribute toward the supply of a public good in contrast to refraining from excessive use of a CPR, for instance, the challenge is to induce individual participants to take action rather than to persuade them to avoid acting in a manner that is detrimental to social welfare.[5] Monitoring is seldom a problem when it comes to identifying contributions to the supply of a public good (e.g. improved knowledge of the functioning of key ecosystems); contributors will be anxious to make their contributions known to others. Nor is there any need to worry about social boundaries. Because (pure) public goods - unlike CPRs - are not subtractable or rival, additional members can join the group with no negative consequences in terms of benefits accruing to the original members. To the extent that they are prepared to join in the group's cost-sharing mechanism, in fact, the contributions of new members will ease the burdens on the original members of the group. On the other hand, it is only fair to acknowledge that eliminating free riders in the context of public-goods problems (i.e. actors seeking to enjoy the benefits of a public good while contributing nothing to its supply) may well require the establishment of effective compliance mechanisms. As experience with efforts to implement many international environmental regimes makes clear, it is no easy task to

[5] . Whereas CPRs are goods that are non-excludable and rival or subtractable, (pure) public goods are both non-excludable and non-rival (Olson 1965).

induce individual actors to live up to the financial commitments they make at the outset, even when there is no difficulty in monitoring the extent to which they make contributions to the supply of a public good.[6] In some cases, parties cannot even bring themselves to file routine status reports in a timely manner.

Another common problem type arises in highly asymmetrical situations where the actions of one party (e.g. emissions of airborne pollutants, diversions of upstream waters into irrigation canals) produce externalities detrimental to the welfare of one or more of its neighbors. The problem here is to find ways to induce those who are sources of negative externalities to pay attention to the effects of their actions on the welfare of others. Although sanctions may be relevant in such settings, monitoring is seldom a problem since victims have well-developed incentives to prove that harm is being done to them and to provide evidence that will satisfy unbiased observers. Even more relevant is the observation that parties involved in such asymmetrical situations may well experience incentives to expand the set of issues included in problem sets or in negotiations aimed at resolving their differences (Sebenius 1983). If this is done in such a way that asymmetries relating to individual issues included in a larger package are matched with one another so that the overall bundle of issues under consideration becomes (roughly) symmetrical, what was initially a more or less severe conflict of interest can be transformed into a coordination problem in the sense that the benefits flowing to each of the participants are sufficiently large to suppress any incentives they may experience to cheat on the terms of their agreement.

---

[6] . In some cases, an individual actor may place such a high value on a public good that it is willing to shoulder the cost of supplying the good without exerting pressure on other members of the group to make contributions. In the literature on public goods, such a group is known as a privileged group (Olson

As this last observation suggests, the character of environmental problems is not wholly determined by nature. There is often room to influence the way a problem is framed for purposes of regime formation or institution building. This is especially true during the stage of agenda formation when the issues under discussion have not been crystallized through lengthy bargaining processes (Kingdon 1995, Young 1998). In some cases, this is a matter of spelling out the character of relatively specific issues (e.g. how should we think about the jurisdiction of coastal states in adjacent marine areas?). More often than not, however, the problem involves highly complex issues (e.g. should we include the full range of greenhouse gases within the framework of the climate regime) or even the pros and cons of joining together several complex issues (e.g. ozone depletion, climate change, long-range air pollution) for purposes of regime building. It does not require any wholesale adoption of the tenets of social constructivism to see that there is often considerable room for the development of different ways to think about complex environmental problems (Wendt 1992). While the views of individual actors regarding such matters are often driven by prevailing conceptions of their own interests, there is substantial evidence that cognitive forces that are not easy to reduce to simple expressions of actor interests play a role in defining issues and developing the discourses in terms of which they are addressed (Litfin 1994). This may lead to convergent characterizations of individual problems as cases of one or, at most, a few problem types. But the evidence arising from actual social practices does not warrant any simple conclusion in this regard.

1965). Among students of international relations, an actor that shoulders the full burden of supplying a public good is known as a hegemon (Keohane 1984, Snidal 1985).

It is worth noting as well that efforts to determine what problem type a specific environmental concern exemplifies will often be afflicted by uncertainty either because the biogeophysical mechanisms at work are poorly understood or because there are disagreements regarding either the extent to which the problem is anthropogenic in origin or the nature of the socioeconomic forces giving rise to the problem. In such cases, there are apt to be sharp disagreements about the characterization of problems coupled with a pronounced tendency to come to terms with uncertainty by selecting characterizations as much on the basis of self-interest as on the basis of any objective or unbiased interpretation of available evidence. Settings of this sort often give rise to strong temptations to eliminate uncertainty by seeking to fit the full spectrum of environmental problems into a few familiar categories. Some recent writings on international environmental problems, for example, advance the claim that all such concerns are ultimately CPR problems, so that the challenge in each and every case is a matter of devising suitable arrangements to avoid the tragedy of the commons (Barkin and Shambaugh 1999). Comforting as this approach to uncertainty may be in analytic terms, however, the evidence pertaining to actual cases does not support such a claim. Problems vary; opportunities for devising alternative characterizations of individual problems are substantial, and negotiations frequently begin in settings where there is no clear consensus among the participants regarding the proper characterization of the problem.

The conclusions I draw from this account are that the larger universe of issues associated with environmental changes encompasses a range of problem types that are not reducible to a single homogeneous set and that the characterization of individual cases is by no means a routine – much less objective – process. Under the circumstances, there is no reason to expect a single set of design principles stated in the form of necessary conditions for achieving robustness or sustainability to stand up to vigorous testing. One size does not fit all when it comes to the creation of effective environmental regimes; design principles derived from a study of some members of the larger universe of environmental problems run the risk of failing to produce the desired outcomes or leading to highly inefficient results when applied to others.

Does this mean that the search for design principles is a dead end and ought to be abandoned? Although the pitfalls associated with this approach are obviously severe, I believe it would be premature to reach such a conclusion. The development of design principles stated in the form of necessary conditions for success in problem solving is an attractive prospect, and we cannot rule out the possibility that additional research will turn up a set of conditions of this sort that hold across the entire universe of environmental problems or even sizable subsets of the universe. Even so, it seems clear that we will not be able to count on the design principles approach during the foreseeable future as a comprehensive method for bringing knowledge to bear on solving a variety of environmental problems ranging from the depletion of local fish stocks to climate change or the loss of biological diversity. Does the argument of the previous section, then, imply that we should embrace the opposite view, asserting that each environmental problem is unique in a manner that leaves us with no alternative to treating every case as a universe of one? Although the reasoning underlying this reaction is easy enough to comprehend, I do not find such a conclusion any more satisfactory as a basis on which to initiate a constructive dialogue between analysts and practitioners who share an interest in solving a variety of more or less pressing environmental problems. What is needed, at least at this stage, is an intermediate approach, one that avoids both the pitfalls of excessive generalization and the limitations arising from the treatment of each environmental problem as a unique case.

2. Institutional Diagnostics

The alternative I propose and consider in this section may be described as institutional diagnostics. The defining feature of the diagnostic approach is an effort to identify important features of issues arising from environmental changes that can be understood as diagnostic conditions coupled with an analysis of the design implications of each of these conditions. One useful way to think about this approach is to treat it as an exercise in mid-range generalization coupled with a liberal use of ceteris paribus assumptions. Instead of attempting to identify any given problem as a case of some generic problem type (e.g. a CPR) and then seeking to apply conditions regarded as necessary to a successful treatment of all instances of that type, the diagnostic approach seeks to disaggregate environmental issues, identifying elements of individual problems that are significant from a problem-solving perspective and reaching conclusions about the design features needed to address each of the elements identified.[7] This approach does not yield design principles or, for that matter, any generalizations that apply across the entire universe of environmental problems. Rather, it establishes a procedure in which environmental problems are considered on a case-by-case basis and prescriptions or recommendations are developed that take into account the particular combinations of conditions exhibited by individual cases. Taken together, the combination of conditions occurring in any particular case may be unique. But so long as the individual elements of the set are recognizable conditions that have identifiable implications for the design of management regimes, it is

[7] . Construed in this way, the practice of institutional diagnostics has something in common with Elster's account of mechanisms (Elster 1999).

possible to formulate recommendations for specific cases that are based on the application of mid-range generalizations.[8]

Beyond this, we must recognize at once a distinction between simple diagnostics and complex diagnostics. Simple diagnostics involves a process in which the diagnostic conditions associated with a specific environmental problem are examined one at a time with the objective of determining their individual design implications. The assumption implicit in this procedure is that the individual diagnostic conditions do not interact with one another to any significant degree. This assumption may be appropriate in some cases, at least as a first approximation or as a rough-and-ready guide. When this is the case, simple diagnostics can prove sufficient as a procedure for formulating recommendations pertaining to institutional design. But it is easy to see that there are also cases in which individual diagnostic conditions will interact with one another in ways that are too important to ignore. When interactions of this sort occur, it is essential to supplement simple diagnostics with some form of complex diagnostics.

2.1 Simple Diagnostics

The core of simple diagnostics is a set of three linked procedures: identifying a range of diagnostic conditions, evaluating the design implications associated with each of these conditions, and developing the interpretive skills needed to apply these propositions to specific cases. Informal efforts to make use of procedures of this sort have long been a feature of exercises in regime formation at all levels of social organization. But the idea of developing institutional diagnostics into a self-conscious and more systematic approach to the design of

---

[8] . Note that this approach bears a distinct resemblance to the procedure that an ecologist asked to make recommendations for the management of a specific ecosystem (e.g. a woodlot) would employ.

environmental or resource regimes is largely unfamiliar. There is little prospect that we can devise a typology of diagnostic conditions that is exhaustive. As in most other diagnostic endeavors, the possibility of encountering conditions associated with specific environmental problems that are unfamiliar but turn out to have important prescriptive implications is always present. Even at this early stage, however, it is possible to give a substantial account of simple institutional diagnostics treated as a procedure for addressing the institutional dimensions of environmental change.

Despite the problems facing efforts to construct typologies, it is easy to identify a number of distinct diagnostic conditions that have significant implications for the (re)design of institutions. Thus, Table 2 encompasses three major classes of diagnostic conditions: ecosystem properties, actor attributes, and implementation issues. Ecosystem properties are features of relevant biogeophysical systems (and of our knowledge about them) that have important consequences for institutional design. Actor attributes are characteristics of the set of actors whose behavior gives rise to environmental problems which need to be taken into account in designing arrangements to solve or ameliorate these problems. Implementation issues are considerations relating to the fulfillment of institutional commitments that are important to the performance of arrangements created to address specific environmental problems. A brief discussion of each of these classes of conditions will serve to flesh out the idea of simple diagnostics.

TABLE 2 SIMPLE INSTITUTIONAL DIAGNOSTICS

Diagnostic Condition	Design Implication
A. Ecosystem Properties	
Nonlinear or chaotic systems/ surprises	*Early warning devices/ rapid response capability*
Problem duration	*Management structures*
Functional interplay	*Coordination Mechanisms*
Uncertainty/imperfect knowledge	*Social learning/adaptability*
B. Actor Attributes	
Variability of members' political/economic systems	*Flexibility*
Heterogeneity of member interests	*Issue kinkages*
Asymmetries in causal responsibility	*Emphasis on equity*
Asymmetries in capacity	*Capacity building/ technology transfers*
C. Implementation Issues	
Violation tolerance	*SIRs*
Incentives to cheat	*Role of sanctions*
Behavioral transparency	*Monitoring procedures*
Malleability of rules	*Certification/deterrence*

Some systems are chaotic in the sense that they are subject to nonlinear and often poorly understood changes that can occur suddenly and produce major surprises (Ludwig et al. 1993). Those who regard the problem of climate change as particularly serious, for instance, typically assume that the Earth's climate system is chaotic in this sense and point to indications in the paleoclimatic record of remarkably rapid and far-reaching non-anthropogenic changes in this system occurring at various times in the past. The institutional implications of this condition are straightforward. To the extent that the relevant biogeophysical systems are subject to chaotic behavior it is important to devise effective early warning systems and to build rapid response capabilities into the design of institutional arrangements. The significance of this finding is particularly important in settings, such as contemporary international society, where prevailing constitutive arrangements tend to impede efforts to respond quickly and decisively to changing circumstances.

The design implications of problem duration and functional interdependence, on the other hand, are quite different. While some problems can be solved once and for all, others involve ongoing activities that are likely to continue indefinitely. It is probable, for example, that the problem of ozone depletion will be solved by phasing out the production and consumption of ozone-depleting substances, with the result that the ozone regime will work its way out of a job.[9] By contrast, any regime designed to govern human harvesting of renewable resources (e.g. fish or forests) must be based on the expectation that its services will be needed indefinitely. This distinction yields the following design implication. The longer the duration of the problem, the more sense it makes to invest resources in the development of sophisticated administrative arrangements, funding mechanisms, and dispute-resolution procedures. Ad hoc arrangements that seem perfectly adequate to handle short-term needs will seem seriously deficient as

---

[9] . The impacts of ozone depletion will be felt for some time to come due to the fact that some ozone-depleting substances have lengthy periods of residence in the Earth's atmosphere. Once satisfactory alternatives are introduced, however, it is unlikely that producers and consumers will revert to the use of ozone-depleting substances.

permanent arrangements. For its part, functional interdependence is a matter of the extent to which the problems dealt with under the terms of separate regimes interact with one another. There is no need to be concerned about links between regimes in cases, such as the management of fish stocks in the Bering Sea and the Barents Sea, where the relevant biogeophysical systems are unrelated. But in cases like ozone depletion and climate change, where the functional interdependencies among distinct problems are substantial, it would be foolish not to think carefully about the establishment of institutional links or coordination mechanisms (including nesting and agreed-upon divisions of labor) in building regimes to deal with environmental problems (Aggarwal 1998). In extreme cases, it may make sense to consider merging distinct institutional arrangements in the interests of coming to terms with the implications of functional linterdependencies.

A related, though somewhat different, diagnostic condition involves the state of knowledge about ecosystem properties. In some cases, knowledge of the relevant systems is highly sophisticated and the likelihood of surprises is low, although recent experience with efforts to manage marine fisheries should warn us against any tendency to become complacent in this regard (Dobbs 2000). But in other cases, such as climate change, it is apparent that the biogeophysical systems are unusually complex, understanding of the behavior of these systems is limited, and the probability that unexpected but highly significant events will occurr is high (Bolin 1997). Of course, one response to this condition is to allocate substantial resources to improving knowledge regarding the behavior of the relevant systems. But this response is likely to yield long-term improvement in management capabilities at best. In the meantime, one significant design implication arising from this condition features the adoption of a precautionary approach and, as a result, the establishment of relatively large margins of safety to avoid triggering unintended but irreversible changes in important systems.

Turning to actor attributes, another set of diagnostic conditions and design implications comes into focus. When the actors involved in a problem – whether they are nation states or various types of nonstate actors – exhibit a high degree of heterogeneity with regard to their internal economic and political systems, it will make sense to create regimes that couple the rules and procedures needed to solve problems with a high degree of latitude concerning the procedures individual members are authorized to use in implementing these arrangements. Efforts to guide the behavior of rational, self-interested individuals through appeals to legitimate authority in contrast to measures that alter incentives are unlikely to succeed. But by the same token, incentive mechanisms may have little impact on the actions of those whose behavior is governed by the logic of appropriateness (March and Olsen 1998). Consider the problem of climate change as a case in point. A successful regime in this area must be able to guide the actions of more the 150-200 countries that vary greatly in terms of the character of their political systems, the nature of their economic systems, and their stage of development. As a result, any effort to mandate uniformity with regard to the policy instruments that member states employ to implement the terms of the climate regime within their domestic jurisdictions will almost certainly end in failure.

Heterogeneity with respect to the interests of relevant actors, on the other hand, is another matter. Some observers believe that heterogeneity of this sort is a recipe for conflict that will make problems more difficult to solve. But this is not always the case. In some situations, heterogeneity regarding actor interests actually increases the scope for striking institutional bargains that satisfy the essential interests of all parties concerned. The fundamental design implication of this condition concerns the benefits of expanding the scope of institutional arrangements by adding issues and opening up opportunities to devise package deals in which individual actors make concessions on issues that are of lesser importance to them in return for concessions on the part of others regarding their priority concerns (Sebenius 1983). This process is well-known in domestic settings where it is often described as logrolling or legislative bargaining and sometimes produces packages of institutional arrangements that are more responsive to the parochial interests of individual actors than to the pursuit of common or societal goals like sustainability. The process is less familiar at the international level where individual issues are normally dealt with in separate forums. But as the recent effort to devise a package of provisions relating to climate change that encompasses targets and timetables on the one hand and policy instruments including joint implementation and emissions trading on the other makes clear, there are opportunities for devising package deals at the international level as well.

Asymmetries in causal responsibility and variations in capacity to implement institutional arrangements are additional conditions involving actor attributes that have significant design implications. When some actors bear the major burden of responsibility for causing environmental problems while others are likely to be most significantly affected by their consequences, success in creating effective institutions will turn on issues relating to equity. Here again the case of climate change is instructive. So long as the United States - currently the source of approximately a quarter of all greenhouse gas emissions - is either unable or unwilling to accept responsibility for its role in climate change, especially with regard to potential impacts on low-lying or small-island states which could be damaged severely by rising sea levels, there is little prospect of creating an effective climate regime. One interesting response to this condition centers on asymmetries in the capacity of members countries to handle the burden of coping with climate change. Thus, the United States is not only the largest source of greenhouse gas emissions, it also has the greatest capacity of any of the members of international society to pay for various forms of mitigation and adaptation. One way to address the equity problem arising from asymmetries in causal responsibility, then, is to devise an arrangement under which those with the greatest capacity to address the problem help those who are less fortunate through the development of mechanisms featuring technology transfers and capacity building. The rationale underlying the Clean Development Mechanism (CDM) in the case of climate change reflects this way of thinking. Whether or not this approach proves successful in the case of climate change, the generic design implications are clear. Asymmetries in causal responsibility and in ability to cope with problems must be recognized in the design of arrangements created to address the relevant problems,

Diagnostic conditions involving implementation issues are somewhat different in character from ecosystem properties and actor attributes. They deal with matters of institutional performance in contrast to features of the relevant biogeophysical systems or attributes of the sets of actors involved. But these concerns can be approached in terms of institutional diagnostics as well. There is considerable variance among regimes in terms of both violation tolerance and incentives to cheat. Some regimes (e.g. many arms control arrangements) will unravel or collapse in the face of any serious violations or even unsubstantiated claims regarding violations. Others can withstand relatively frequent violations without having their effectiveness called into question. In the case of climate change, for instance, it is aggregate trends in levels of greenhouse gas emissions that count rather than the compliance of individual actors under specific circumstances. What this means is that the lower the level of violation tolerance associated with a regime, the greater will be the need for systems of implementation review (SIRs) capable of monitoring the behavior of individual members in some detail (Victor, Raustiala, and Skolnikoff 1998). Similar observations are in order regarding incentives to cheat. It is unrealistic to hope that most major environmental concerns will take the form of coordination problems (Stein 1982). Yet there is substantial variance among problems with regard to incentives to cheat and especially the strength of incentives to cheat relative to the gains derived from membership in the regime. Where incentives to cheat are comparatively weak, there will be no need to devote a great deal of time and energy to the development of credible sanctioning procedures. The stronger the incentives to cheat, on the other hand, the greater the importance of sanctions that individual actors find credible (Downs, Rocke, and Barsoom 1996).

Similar comments apply to behavioral transparency and rule type. The degree to which the behavior of relevant actors is transparent is subject both to natural variation and to variation attributable to regime design. In some cases (e.g. the clearcutting of forests), it is difficult to engage in behavior leading to environmental problems in a covert or clandestine manner. Monitoring is not a major problem in such situations. More interesting, however, are cases in which behavioral transparency is naturally low but subject to manipulation on the part of those in a position to influence institutional design. Consider cases like intentional oil pollution at sea and ozone depletion in these terms. Behavioral transparency in the case of oil pollution increased dramatically following a switch from rules based on discharge standards to rules based on equipment standards (Mitchell 1994). In the case of ozone depletion, it quickly became apparent that transparency would be much higher with regard to the actions of a handful of producers in contrast to millions of consumers, even though the consumption of ozone-depleting substances is the ultimate concern (Benedick 1998). This suggests not only that the importance of investing in monitoring systems is a function of the degree to which relevant behavior is naturally transparent but also that there is often considerable scope for designing rules in such a way as to increase transparency. In extreme cases, rules can be framed in a manner that places the burden of proof wholly on the subject. That is, subjects may be required to prove compliance rather than being treated as compliant unless and until some credible evidence of violations comes to hand. Short of this, however, there are often many different ways to frame the principal rules of regimes designed to deal with environmental problems (Young 1999b). The value of investing resources in the framing of rules will vary as a function of other compliance concerns, including violation tolerance and incentives to cheat. But when rules are malleable, there is much to be said for thinking carefully about the pros and cons of different ways of framing them.

What can we conclude from this account of simple diagnostics? The basic logic of this procedure is clear. By examining diagnostic conditions one at a time, it is possible to think clearly about their implications for institutional design. There is considerable variance among actual environmental issues in terms of the extent to which they exhibit these conditions. As a result, we cannot expect this diagnostic approach to yield generalizations of the sort associated with the idea of design principles. What is more, this diagnostic procedure resembles diagnostic procedures in other fields in the sense that using it to reach conclusions about specific problems (e.g. ozone depletion, climate change) requires considerable skill. It is perhaps inappropriate or unfair to describe diagnostics as an art rather than a science. But it is clear that those who use such procedures vary greatly in terms of their ability to arrive at accurate results. Whereas the unskilled diagnostician proceeds to examine relevant conditions in a mechanical fashion, the truly talented diagnostician acquires an almost intuitive feel for the significance of key diagnostic conditions in specific situations. Of course, the importance of this skill depends upon the complexity of the situation at hand. Almost any well-trained person can diagnose problems that are simple and straightforward. Real skill is required to deal with complex problems. This is one reason why climate change is so intriguing to students of environmental change. Not only is climate change likely to have profound consequences for human welfare, but this problem is perhaps the most complex environmental issue we have faced to date from the perspective of institutional diagnostics.

2.2 Complex Diagnostics

The idea of simple diagnostics is appealing precisely because it offers a procedure for thinking about the links between analysis and praxis with regard to institutional design that is relatively easy to understand. Yet this procedure is based on the powerful assumption that individual diagnostic conditions do not interact with one another in significant ways. This assumption may be perfectly reasonable under some circumstances. But it is clear that interactions among relevant conditions will be too important to ignore in other cases. When this happens, there is no alternative to adopting the procedure of supplementing simple diagnostics with a consideration of complex diagnostics. The range of interactions among two or more diagnostic conditions is great; a useful typology of these interactions is currently beyond our grasp. Nonetheless, it is not hard to point to concrete cases that provide clear illustrations of the idea of interactions between diagnostic conditions.

One prominent example centers on interactions between nonlinear or chaotic systems and uncertainty or imperfect knowledge regarding the behavior of such systems. The origins of these interactions are apparent. Nonlinear systems, especially those in which surprises constitute a common occurrence and changes of state can occur quite suddenly, are far more difficult to understand than systems featuring stable equilibria. Not only are the systems themselves unusually complex, the analytic tools needed to understand them are also comparatively weak. An obvious case in point is the Earth's climate system. There are good reasons to believe that this poorly understood system is subject to nonlinear and sudden, surprising changes that may have profound consequences for human welfare. Yet our capacity to forecast - much less to predict – the occurrence of these changes is limited. What does this mean for the design of the climate regime? One response is to downplay the problem on the grounds that most societies have more pressing issues to contend with and that humans will learn their way out of this problem as they have done with others. But an alternative response that may prove more convincing is to apply the precautionary principle which, in this context, would mean including a margin of safety in setting targets and timetables within the climate regime. The argument here rests on the usual case for insurance; it makes sense to pay a small – albeit significant – price now to minimize the impact of potentially catastrophic occurrences at some later time.

A different sort of illustration involves interactions between problem duration and efforts to come to terms with issues like incentives to cheat and behavioral transparency. In dealing with problems that can be solved within a relatively short time frame, it is appropriate to limit investments in compliance mechanisms (e.g. programs designed to socialize key actors over the long term) that are costly and make sense only when they can be amortized over a long period of time. Thus, it seems reasonable to draw a distinction in these terms between the problem of ozone depletion on the one hand and the problems of climate change and the loss of biological diversity on the other. Handled properly, phasing out the production and consumption of ozone-depleting substances will be complete within a decade or two at the most. Once solved, it is not likely that this problem will reemerge in another form.[10] As a result, it is sensible to devise ad hoc arrangements, such as the Montreal Protocol Multilateral Fund, to persuade actors to comply with the rules pertaining to the production and consumption of ozone-depleting substances and to avoid creating costly arrangements that may take on lives of their own and stay in business long after the problem is solved. Contrast this situation with the cases of climate change and biological diversity. These problems will be with us over the long term. The struggle to protect biological diversity, in particular, is destined to become a permanent item on the environmental agenda. The implications of this set of circumstances for institutional design are clear. Although the initial costs may be high, it makes sense under these conditions to make substantial investments early on in efforts to alter entrenched rights (e.g. familiar rights of private property owners) and to socialize actors into complying with regulatory rules as a matter of appropriate behavior rather than as a consequence of calculations of benefits and costs.[11]

---

[10] . Of course, it is possible - even probable - that problems analogous to ozone depletion in the sense that they are externalities of the use of chemicals developed for other purposes will arise from time to time.
[11] . This argument may go some way toward justifying the expenditure of more time and energy on the organization of the Global Environment Facility (GEF) than the Multilateral Fund.

When rules are malleable but violation tolerance is low and incentives to cheat constitute a real concern, another pattern of interaction surfaces. Different ways of framing key rules can affect the prospects for eliciting compliance, in some cases dramatically. Where the rules require large numbers of actors to refrain from behavior that is not naturally transparent, for instance, either the costs of obtaining compliance or the incidence of violations will rise sharply. Imagine the difficulties associated with persuading or compelling millions of consumers to comply with rules regarding the consumption of products manufactured through processes involving substantial emissions of greenhouse gases. Eco-labeling will lead some consumers to change their behavior at the margin in this connection. But such initiatives are hardly likely to solve the problem. In a case like this, there is much to be said for directing attention to alternative ways of framing the rules. One interesting option is to shift the burden of compliance from large numbers of consumers to much smaller numbers of producers. Requiring automobile manufacturers to meet well-defined fuel economy standards is a familiar case in point. A complementary strategy is to place the burden of demonstrating compliance on subjects rather than assuming that the actions of subjects conform to the rules unless and until proven to be non-compliant. These are strong measures, and there are obviously good reasons to proceed with care in the framing of rules in specific cases. But it is clear that when concerns about violation tolerance and incentives to cheat are substantial yet rules are malleable, those concerned with problem solving will want to pay close attention to the pros and cons of alternative ways of formulating the rules.

These intuitively appealing examples can do no more than illustrate the idea of complex diagnostics. The range of potential interactions, including interactions involving three or more diagnostic conditions, is obviously great. In analytic terms, this leads to an interesting tradeoff. The procedures associated with simple diagnostics are much easier to use. But they may yield only prescriptions that are either obvious even to those who do not use systematic diagnostic procedures or inappropriate. Efforts to engage in complex diagnostics, by contrast, are more likely to yield counterintuitive results, but they may also eventuate in convoluted assessments whose design implications are both vague and difficult to explain. Fortunately, there is no need to make any definitive choice between simple and complex diagnostics. Of course, the fact that time and energy are finite will force analysts and practitioners alike to make tradeoffs in this realm in dealing with specific cases. But there are no conceptual or analytic barriers to the adoption of a mixed strategy that features the use of both forms of institutional diagnostics at the same time.

2.3 Common Diagnostic Errors

Like most other diagnostic endeavors, the practice of institutional diagnostics may yield erroneous conclusions in specific cases. Partly, this is a matter of the skill that the diagnostician brings to the task. It does not require any great insight to recognize that there is considerable variation among practitioners with regard to their diagnostic skills. In part, however, diagnostic failures are products of standard errors that can afflict the efforts of even the most skillful diagnostician. We are not now in a position to provide a comprehensive list of these common errors; it may never be possible to produce a definitive list. But consider the following common problems as illustrations of this concern: missing data, inappropriate models, unidentified interactions, and negotiated assessments.

Missing data constitutes an obvious limitation, especially in efforts to assess ecosystem properties. Are the relevant systems prone to nonlinear changes? If so, are shifts from one state to another likely to occur suddenly as envisioned in the idea of punctuated equilibria or more gradually as envisioned in most discussions of evolutionary change? In situations where there are large numbers of cases that can be examined empirically or even subjected to controlled experiments, it is comparatively easy to answer questions of this sort. But in cases, such as the Earth's climate system, involving biogeophysical systems that are large, complex, and unique, the data needed to evaluate important diagnostic conditions are apt to be few and far between. Efforts to determine whether the trends of the last few decades in global mean temperatures at the Earth's surface are merely short-term cycles or early stages in a longer-term pattern of global warming regularly run into problems of this sort. In the absence of a well-tested, predictive model of climate change, the natural reaction is to compare recent trends with a number of past cases of climate change. Although such efforts have yielded some suggestive results, the sparseness of data on past cycles places severe limits on diagnostic efforts of this sort.

A somewhat more insidious diagnostic error arises from the influence of models or analytic constructs that acquire a high level of credence but that ultimately prove to be inappropriate or applicable only to a limited set of cases. A classic example of this problem involves the development and application of single-species, logistic models to assess the status of fish stocks and to prescribe regulatory procedures intended to produce maximum sustained yields (MSY) in individual fisheries over indefinite periods of time (Larkin 1977). In fact, large marine ecosystems feature complex interactions among different species and regularly undergo nonlinear changes following significant disturbances, such as the depletion of species that make up sizable segments of the biomass (Sherman 1992, Wilson et al. 1994, NRC 1996). The use of MSY models to diagnose the condition of stocks and to make decisions about allowable harvest levels under such conditions frequently leads to undesirable and, in some cases, catastrophic results (McGoodwin 1990). Yet the use of such models dies hard. Not only are analysts apt to hold onto simple models of this sort unless and until better alternatives become available, but also such models often make their way into the received and largely unquestioned wisdom of the relevant community. As the case of marine fisheries makes clear, this can lead to a disturbing record of erroneous diagnoses even among otherwise highly sophisticated observers.

Similar remarks are in order regarding unidentified interactions. There is a natural tendency to think about environmental problems in terms of events occurring within stand-alone systems. Given the complexity of the dynamics of most human-dominated ecosystems, this tendency is certainly understandable. Yet it is increasingly clear that interactions among these systems are often significant and sometimes critical. Both ozone-depleting substances and many substitutes for these substances are greenhouse gases (Oberthür 1999). The fate of the Earth’s forests will have critical consequences both for climate change and the maintenance of biological diversity. Shifts in the composition of species in large biogeophysical systems can accelerate or retard processes of desertification or soil erosion. Any successful effort to cope with environmental problems must rest on assumptions about the boundaries of relevant systems and the significance of interactions among distinct systems. Appealing as it is in conceptual terms, the familiar ecological principle that everything is related to everything else (Commoner 1972) does not offer much practical guidance for those seeking to solve specific problems. Yet a failure to take into consideration the consequences of functional interdependencies is a common source of diagnostic errors.

Beyond this lies the problem of negotiated assessments. In the absence of decisive evidence regarding the status of biogeophyiscal systems and the impacts of anthropogenic drivers, diagnoses are often subject to political processes. Even the periodic assessments provided by the Intergovernmental Panel on Climate Change (IPCC), a body that has been praised deservedly for the rigor and unbiased character of its procedures, are ultimately negotiated in a setting that is by no means strictly scientific (Edwards and Schneider forthcoming). In one sense, there is no escape from such processes. All responses to environmental problems – including decisions to do nothing - rest on assessments regarding the behavior of the relevant systems. All assessments, regardless of the care with which they are carried out, involve judgments that can never be strictly objective or unaffected by the beliefs and values of those who make them (Miller and Edwards forthcoming). Even so, there is a striking difference between assessments that are based on careful and transparent procedures and those that are products of obscure bargaining among obviously self-interested actors. The contrast between the work of the IPCC or the International Council for the Exploration of the Sea (ICES) and the activities of the scientific committees attached to many regimes dealing with marine living resources in these terms is striking. Institutional diagnoses cannot escape the need for judgment. The important point is to avoid setting too much store by negotiated assessments.

In the nature of things, institutional diagnosis is a fallible procedure. The role of judgment is substantial, and the absence of simple formulas places a premium on the contributions of the skilled diagnostician. There is no surefire way to avoid all diagnostic errors. But one way to minimize the incidence and impact of such pitfalls is to identify the major categories of diagnostic errors as clearly as possible and to encourage those engaged in institutional diagnostics to consider the relevant dangers carefully in dealing with specific cases. An awareness of the shortcomings of MSY models, for instance, does not guarantee success in designing successful regimes for marine living resources. But it may well play some role in allowing those responsible for the creation and implementation of fisheries regimes to avoid pitfalls that have occurred again and again in efforts to manage fisheries or, for that matter, many other human-dominated ecosystems (Vitousek et al. 1997).

3. Final Words

Some observers take the view that it is desirable to separate analysis and praxis in thinking about the institutional dimensions of environmental problems. The argument here turns on the proposition that it is desirable to insulate analysis from the more political processes involved in efforts to solve specific problems. The principal concern, on this account, is that interactions between analysts and practitioners will corrupt efforts to produce usable knowledge and, at the end of the day, lead to results that are both unsatisfactory in analytic terms and of little value in applied terms. Although this concern is real and understandable, I do not subscribe to the conclusion it suggests. Analytic processes themselves have a political dimension, regardless of our efforts to achieve objectivity. There is no need to adopt all the arguments often grouped under the heading of the social studies of science to acknowledge this fact (Jasanoff and Wynne 1998). But even more important, in my judgment, is the fact that an ongoing dialogue between analysts and practitioners can prove beneficial to both communities, helping analysts to test existing ideas and generate new insights on the one hand and assisting practitioners to broaden their thinking and to avoid simplistic analogies or inappropriate responses to specific problems on the other. Both the development of design principles and the pursuit of institutional diagnostics can serve as vehicles for fostering such a dialogue; there is no need to make a choice between these approaches to the production of usable knowledge. Nonetheless, I am convinced that institutional diagnostics constitutes an approach that is particularly well-suited to facilitating productive exchanges between analysts and practitioners and that deserves increased attention among those desiring to generate usable knowledge pertaining to the institutional dimensions of environmental change.