CHAPTER 2

Institutional Analysis and Common-Pool Resources

The substantive concerns of this book involve understanding how rules affect the behavior and outcomes achieved by individuals using common-pool resources (CPRs). To address these concerns effectively requires us to raise fundamental questions about how to explain observed behavior in laboratory and field settings. The three basic questions that we identified in chapter 1 are theoretical questions. They have to do with how we think about CPRs. Policies are fashioned from the way that public officials, citizens, and scholars think about problems. We hope that our readers recognize how important the ideas used in policy analysis are. Unless empirically well-grounded theories are developed to enhance the prediction and understanding of behavior, the likelihood of changing rules so as to improve outcomes is slim.

Models, Theories, and Frameworks

Empirically well-grounded theories, however, do not just appear out of thin air. Nor is all theoretical work accomplished at the same conceptual level. Theoretical work related to the study of rules proceeds on at least three different levels: formal models, theories, and frameworks. These levels are all important in the long-term development of empirically grounded theory.

The Formal Model Level

Formal models make explicit assumptions about the elements and structure of a particular situation and use the logical tools of a theory to derive predictions about the likely outcomes of a particular set of parameters. Chapters 3–6 of this book are examples of work at this conceptual level. In chapter 3, we initially develop the simplest possible two-person models of four types of CPR games—appropriation, assignment, provision, and monitoring. In chapter 4, we focus on how various types of rule configurations affect the structure (and thus the predicted outcomes) of the assignment game that we initially
present in chapter 3. In chapter 5, we expand the two-person appropriation game into an \(N\)-person game that can be tested in an experimental laboratory. Chapter 6 extends the analysis to a time-dependent situation.

Models are always models of something else. Our formal models are models of CPR situations that draw on a more general theory—noncooperative game theory—in providing the logical tools and techniques for building specific models. Game theory is at a more general conceptual level than the models that apply game theory to particular problems. \(^1\) Further, one can generate many game-theoretical models of similar situations. For example, the model of a two-person appropriation game using complete information, presented in chapter 3, could be contrasted with a two-person appropriation game using incomplete information. This would be a different model with the potential for a different predicted equilibrium. Without going outside the confines of game theory, alternative models of the same situation can be developed so that the precise implications of using one or another assumption can be explored. Game theory, then, is the metalanguage for game-theoretical models.

The Theory Level

At this conceptual level, theorists are concerned with puzzles that apply to general classes of models rather than specific models. A theory provides a metatheoretical language for formulating, postulating, predicting, evaluating, and changing various models of that theory. A recent concern among game theorists is the number of game-theoretic models that discover multiple, rather than single, equilibria. Theories of equilibrium selection (Harsanyi and Selten 1988) and equilibrium refinements (van Damme 1987) focus on games that have more than one equilibrium and address how the theorist should proceed. Applications of game theory draw on the developments made at the theory level.

In addition to noncooperative game theory, there are many other theories of human behavior upon which scholars interested in CPRs and in institutions can draw. Theories of bounded rationality are quite relevant—as we discuss later—for the analysis of more complex situations that exist in CPR settings. Cooperative game theory is relevant when the players may freely communicate and commit to binding agreements enforced by a third party, such as some forms of bargaining. Where individuals making economic decisions do not encounter strategic interaction, the microeconomic theory of perfect competition is appropriate. To think about, develop, and evaluate diverse theories, one needs a general framework. Before microeconomic theory was fully

---

\(^1\) Duncan Snidal 1985 stresses the important difference between game theory and game models.
developed, the general theoretical framework of classical economics provided the paradigmatic foundations for theoretical work in economics. A framework provides a metatheoretical language for thinking about diverse theories and their potential usefulness in addressing important questions of relevance to the analyst.

The Framework Level

At the conceptual level of a framework, theorists identify the broad working parts and their posited relationships that are used in an entire approach to a set of questions. Frameworks help to organize diagnostic and prescriptive inquiry. The framework we use is called the Institutional Analysis and Development (IAD) framework, which has been the object of considerable thought and reflection by many colleagues over the years. We use the IAD framework as a general organizing tool that helps us develop a long-term research program not only for research on CPRs but also on other problems where individuals find themselves in repetitive situations affected by a combination of factors derived from a physical world, a cultural world, and a set of rules.

Historical Roots of IAD

The IAD framework has its roots in classic political economy (specifically the work of Hobbes, Montesquieu, Hume, Smith, Hamilton, Madison, and Tocqueville); neoclassical microeconomic theory, institutional economics (the work of Commons 1957 and Coase 1937); public choice theory (Buchanan and Tullock 1962; Downs 1957; Olson 1965; Riker 1962); transaction-cost economics (North 1990; Williamson 1975, 1985); and noncooperative game theory (Harsanyi and Selten 1988; Luce and Raiffa 1957; Shubik 1982). The working parts of the IAD framework, which we discuss below, do not always overtly show in an institutional analysis. That is the case with all frameworks. Since a framework orient s the analyst to ask particular questions, it is the questions that are generated by using the framework that appear in most analyses rather than the intellectual scaffolding used by the analyst to diagnose, explain, and prescribe.

The IAD framework has influenced the analysis of a myriad of issues during recent decades. It has been applied to the study of metropolitan organization (Advisory Commission on Intergovernmental Relations 1987, 1988, 1992; V. Ostrom, Tiebout, and Warren 1961; V. Ostrom, Bish, and E. Ostrom 1988); the theory of public goods (V. Ostrom and Ostrom 1977); the suste-

nance of rural infrastructures in developing countries (E. Ostrom, Schroeder, and Wynne 1993); privatization in developed and developing countries (Oakerson et al. 1990); to the study of macropolitical systems (Kaminski 1992; V. Ostrom 1987, 1991; Sawyer 1992; Yang 1987) and to a considerable amount of work on CPR problems (Oakerson 1992; E. Ostrom 1990, 1992; Thomson, Feeny, and Oakerson 1992). Work has been carried on related to patterns of order, not only in the United States but in Bangladesh, Botswana, Cameroon, Ghana, India, Indonesia, Ivory Coast, Liberia, Mali, Madagascar, Nepal, the Netherlands, Nigeria, Norway, Poland, the Sudan, the former Soviet Union, and the former Yugoslavia.

The IAD framework does not limit an analyst to the use of one theory. Depending upon the context of the decision environment, an analyst may in fact use the framework as a foundation for investigating the predictive power of complementary or competing theories and models. The initial research approach we develop in this volume combines the IAD framework with the formal theory of noncooperative games and full rationality. For some field and experimental CPRs, noncooperative game theory is particularly useful, and empirical evidence (at least at the aggregate level) is consistent with predictions. For the experimental settings that we describe in chapters 5 and 6, the suboptimal outcomes associated with the game equilibria of limited-access CPR games are broadly supported by the data. This suggests that in field settings where individuals with short time horizons cannot communicate, do not trust each other, or do not have access to reliable external enforcers, outcomes are likely to be broadly consistent with noncooperative game theory.

In other experimental and field CPR situations described in this volume, the data are not consistent with predictions derived from noncooperative game theory under standard assumptions. In some instances, the changes needed to improve game-theoretical tools for the analysis of CPR problems are relatively minor. As we note in chapter 4, for example, game theory does not distinguish between the types of constraints that affect the structure of a game: the constraints of the physical and biological world and the constraints imposed by the rules that individuals evolve or design to limit what can be done in a particular setting. Since all the rules of the game are considered to be immutable from within the game, the possibility that individuals can themselves change the rules of the game (in a time-out or a different arena) cannot easily be addressed without making the distinctions we introduce in chapter 4.

Closely related to the lack of attention to the distinction between physical and biological constraints and the humanly designed rules of the game is how rules get enforced. An underlying assumption of modern game theory is that the rules of the game are unambiguously enforced by some agency external to the game. How and why agents are motivated to enforce rules fully and fairly cannot, therefore, be addressed, as the enforcers are “outside” the game. To
understand many CPR environments, however, it is necessary to bring the enforcers inside the game. Further, in many CPR field settings, the enforcers are not even different actors but rather the same individuals who appropriate from a resource. In chapter 3, we construct an irrigation game where we allow appropriators in the first position to decide between following or breaking the rules of the game and the appropriators in a second position to decide between monitoring or not monitoring the behavior of the first. To do this we must use a contrivance that a "legal" move within the formal game is to break the rules of the game we are modeling. Without this contrivance, the issue of rule breaking and rule enforcing cannot be addressed by a noncooperative, game-theoretic model. With this contrivance, we are able to show that self-monitoring can lower rule-breaking behavior but never eliminate it. In recent papers, Weisinger and Ostrom (1991, 1993) have shown that external agents cannot fully eliminate rule-breaking behavior either.

We find modern game theory to be a powerful and useful tool for understanding behavior and outcomes within CPR situations, particularly when brought within the umbrella of the IAD framework with the consequent attention paid to rules. However, as we continue to conduct empirical work on CPR situations in field and experimental settings, we have encountered ever greater problems in explaining empirical results with only modest changes in the theoretical tools we use. In chapter 9, we identify experimental findings that cannot be explained relying on small modifications of received theory. These anomalies are closely related to those found in other environments that have led many scholars to challenge theories based on assumptions of complete rationality and unlimited computational capability. Thus, having reached the limits where modern game theory with fully rational players provides consistent theoretical guidance, we apply a theory of bounded rationality to explain the degree of cooperation reached among individuals who are given a chance to devise their own rules.

In all of our work, we have relied on the IAD framework as the general scaffolding that supports our inquiries, helps us identify relevant variables to explore, and provides a broader language that any specific theoretical language we, or other social scientists, might want to use. Consequently, we will provide a brief overview of the IAD framework in this chapter so that we share our general paradigm with others before turning to some of the theoretical tools we use to explore versions of that paradigm.

The Institutional Analysis and Development Framework

Markets, hierarchies, and collective-action situations are sometimes presented as fundamentally different "pure types" of situations. Not only are these situations perceived to be different but each is presumed to require its own
language and explanatory theory. Scholars who attempt to explain behavior within markets may rely exclusively on neoclassical microeconomic theory. Scholars who attempt to explain behavior within hierarchies may rely exclusively on political and sociological theory. Scholars who attempt to explain behavior in a collective-action environment may rely exclusively on noncooperative game theory. Such a view precludes the development and use of a more general explanatory framework that, together with its constituent theories, could help analysts make institutional comparisons and evaluations.

Given the multiple levels of analysis involved in institutional analysis, there are several ways that one can approach a question. One of the first steps that can be taken in an institutional analysis using the IAD approach is the identification of a conceptual unit—called an action arena—that is subsequently the focus of analysis, prediction, and explanation of behavior and outcomes within fixed constraints. Action arenas include an action situation component and an actor component (see fig. 2.1). Action situations refer to the social space where individuals interact, exchange goods and services, engage in appropriation and provision activities, solve problems, or fight (among the many things that individuals do in action situations). In field settings, it is hard to tell where one situation ends and another begins. Life continues in almost a seamless web as fishers move from home to a harbor to a nearby fishing grounds and then to a market where the day's haul is sold.

The observer who wants to analyze the recurrent structure of situations must, however, find ways of separating one situation from another for the purpose of analysis. Further, individuals who participate in many situations must also know the difference among them. The actions that can be taken on the fishing grounds are not the same as those that can be taken in the fish market. An individual who repeatedly is mixed up about what situation he or she is in is not considered to be competent.

What is distinctive about the IAD framework, as contrasted to many frameworks that are closely tied to a single social science discipline, is that all situations are viewed as being composed of the same set of elements. Markets, CPRs, hierarchies, and legislatures are all viewed as being constituted

---

3. An important aspect of the IAD framework is, however, that the analysis of changes in these same parameters is an important part of a full institutional analysis. Thus, for many purposes, we assume a given physical and institutional world and ask what difference these fixed constraints make in outcomes. But, as we discuss later in this chapter and again in chapter 4, many institutional analyses focus precisely on the effects of changing the constraints known as rules. Consequently, one can start an institutional analysis by first looking at the factors that affect action arenas rather than with the arenas themselves.

4. One of the reasons that game theory is particularly compatible with the IAD framework is that it also views all action situations, now conceptualized as games, as constituted of similar working parts.
AN ACTION ARENA IS COMPOSED OF

- An Action Situation involving
  - Participants in
  - Positions who must decide among diverse
  - Actions in light of the
  - Information they possess about how actions are
  - Linked to potential
  - Outcomes and the
  - Costs and Benefits assigned to actions and outcomes

- Actors, the participants in Action Situations who have
  - Preferences,
  - Information-processing capabilities,
  - Selection criteria, and
  - Resources.

Fig. 2.1. Components of action arenas

by a similar set of elemental parts. A minimal action situation is characterized
using seven clusters of variables: (1) participants, (2) positions, (3) actions,
(4) potential outcomes, (5) a function that maps actions into realized out-
comes, (6) information, and (7) the costs and benefits assigned to actions and
outcomes. Since many of these elements are themselves relatively complex,
the variety of action situations that can be constructed from these elements is
immense. Thus at the same time that the framework stresses a universality of
working parts, it also enables theorists to analyze unique combinations of
these universal working parts. Further, each of these parts are constituted by
combinations of physical, cultural, and rule-ordered attributes, as we discuss
later in this chapter.

The Action Situation

We will now discuss the elements of an action situation and then turn to how
actors are conceptualized in this framework.

Participants

The first element of an action situation includes the actors who have become
participants in a situation. This is also the element that links actors, given the
way they are conceptualized, to an action situation. In the minimal action
situation, there is only a single participant. The theories that are relevant to
such a situation include all of the various approaches to decision science,
including linear programming and statistical decision theory. At least two
participants (but only one position) are necessary for an analyst to use game theory. There is a fundamental difference between two-player games and games involving more than two players. The theory of perfect competition (and some voting theories) are limiting cases when the number of players becomes so large that the actions of one player are negligible to others.

Positions

Positions are simply place holders to associate participants with an authorized set of actions (linked to outcomes) in a process. Examples of positions include first movers, bosses, employees, monitors, voters, elected representatives, judges, appropriators, and citizens. In some situations, every participant holds the same position. In others, every participant holds a different position. In most situations, the number of positions is less than the number of participants. The capabilities and limitations of being in a particular position depend on the way the other elements are defined.

Once the other elements of an action situation are specified, for example, a first mover may be a very powerful or a very weak position. That depends upon the options left to others once the person authorized to move first makes a decision. Similarly, being told that a participant is a “boss” does not tell us the full story about the relative status and power of that individual. To get a complete picture, one needs to know more about the actions the individual can take and the outcomes that can be affected. Whether an actor is a “boss” in a civil service system where decisions about hiring and firing are made by others or is the owner of a private business with considerable discretion to hire and fire employees without much need to confer with others affects what it means to be a boss. Even the type of information available to a participant may be tied to the position that actor currently holds. What is essential is that in the IAD framework, analysis is undertaken about the actions that individuals who hold particular positions are likely to take, rather than focusing on individual personalities independent of the structure of the situation in which they are acting. Thus, once all the other components are settled, the full meaning of a position is articulated.

Actions

The third element is the set of actions that participants in particular positions can take at different stages of a process (or, nodes in a decision tree). Examples of actions include decisions to fish or not to fish during a defined time period; to go to one fishing spot or another; and to fight or not with another fisher about fishing in a particular location. These are the actions we examine in the assignment game in chapter 4. In many action situations, the array of
actions that are available is immense and may exceed the capacity of current theoretical tools to analyze. Most analyses attempt to identify only those actions that are the most important in a situation, in the sense that choices made about them make an essential difference in the outcomes achieved.

Potential Outcomes

The fourth element is the outcomes that participants can potentially affect through their actions. Examples of potential outcomes include the quantity of fish caught in a fishing spot, the extent of damage imposed by one participant on another, the physical condition of an irrigation system, or destruction of the regenerative capacity of a CPR. In other words, these are the potential outcomes of individuals interacting with one another in a regularized setting.

Transformation Functions

The fifth element of an action situation is the set of functions that map participants (and/or random actions) at decision nodes into intermediate or final outcomes. In some action situations in economics, these functions are called production functions. They link various combinations of inputs into some type of product. In a voting situation, the transformation function takes the symbolic actions of individuals and produces a collective decision. Transformation functions can be determinate or stochastic in nature. The degree of certainty regarding the transformation function can vary with the situation. In most quid pro quo situations within a defined market, for example, participants know the exact conditions for an exchange to be completed. But in some situations, neither participants nor observers fully understand the complex transformations involved. This is the case, for example, in regard to many fishing grounds where fishery biologists do not yet understand the combination of factors that affect the relationship between fishing effort in one year and availability of fish in the next year.

Information

Closely allied to the type of transformation function is the sixth element—the set of information available to a participant in a position at a stage in a process. When the transformation function is simple and determinant complete information about actions, outcomes, and their linkages may be generated. Many situations generate only incomplete information because of the physical relationships involved or because the rules preclude making all information available.
Payoffs

The seventh element is the set of payoffs that assign benefits and costs to actions and outcomes. Examples include the price of rice offered to the irrigator for crops brought to market, the costs of traveling to a fishing spot, the fines attached to illegal actions, or taxes paid on various activities. Thus, payoffs differ from outcomes as they are the method of assigning positive and negative weights to the outcomes and the actions leading to outcomes. During the monsoon season in an Asian country, particular actions taken in irrigating fields are transformed in a relatively predictable manner into a quantity of rice produced. Thus, the outcomes achieved are consistent from one year to the next. The payoffs achieved, however, may differ radically from one year to the next depending on the costs of inputs (such as labor and fertilizer) and the price that a farmer can command when selling rice. In some formal models, outcomes are not overtly separated from payoffs, but they are implicitly assumed. To understand many situations, however, keeping outcomes and the payoffs assigned to combinations of actions and outcomes is quite essential.

A specification of these seven elements—plus a set of assumptions about the attributes of an actor—is made whenever a theorist undertakes an analysis of a CPR setting, one-shot or repeated. In chapter 3, we provide several examples of how these elements are used to construct different types of CPR games. We consider these to be a minimal set of necessary elements for the construction of theories and models of settings where outcomes depend on the acts of individuals. This is a minimal set in that it is not possible to predict behavior in an interdependent situation without such a specification.

A standard mathematical structure for representing an action situation is a game (Selten 1975; Shubik 1982). The decision environment faced by participants in a well-designed laboratory experiment also represents an action situation. The concept of an action situation is, however, broader than any particular theoretical instance. Any action situation, be it a CPR, a committee, a market, or a hierarchy, can be constructed from these seven elements.\footnote{5. The simplest possible representation of a committee, for example, can be constructed using the following assumptions:}

1. One position exists; that of member.
2. Three participants are members.
3. The set of outcomes that can be affected by the member contains two elements, one of which is designated as the status quo.
4. A member is assigned an action set containing two elements: (a) vote for the status quo and (b) vote for the alternative outcome.
5. If two members vote for the alternative outcome, it is obtained; otherwise, the status quo outcome is obtained.
6. Complete information is available about elements (1) through (5).

For this simplest possible representation of a committee, and using a well-defined model of the
A change in any of these elements produces a different action situation and may lead to very different outcomes. More complex models of CPRs, committees, markets, or other interdependent situations are constructed by adding to the complexity of the elements.\textsuperscript{6}

\textbf{Actors}

To predict how actors will behave, the analyst must make assumptions about four clusters of variables: (1) the preference evaluations that actors assign to potential actions and outcomes; (2) the way actors acquire, process, retain, and use knowledge contingencies and information; (3) the selection criteria actors use for deciding upon a particular course of action; and (4) the resources that an actor brings to a situation. The actor in a situation can be thought of as a single individual or as a group functioning as a corporate actor.

\textbf{Individual Preferences}

In most theories of rational behavior, individuals are presumed able to construct a complete preference ordering over outcomes to which payoffs are assigned. Preference theory is itself a vast subject. Many different theories exist about how actors acquire preferences, what they do when outcomes imply extreme trade-offs among valued objects, and how preferences are assigned when outcomes are unknown. Utility theory is a richly developed body of theory for how individuals assign a valuation—utility—to the outcomes and costs of actions.

\textbf{Individual Information-Processing Capabilities}

To explain how individuals make decisions, the theorist specifies the level of information actors possess and process. A frequent assumption made in theo-
ries of full rationality is that individuals have complete information. Specifying that participants have complete information means that they know

1. the actions that each participant can take at every stage of a decision process and those acts that are governed by a random operator;
2. the intermediate and/or final outcomes that can be reached as a result of the moves of various participants combined with chance moves where relevant; and
3. the preference ranking placed by each participant on all outcomes.

If a participant knows all of the above, the participant knows the full decision or game tree in extensive form. Perfect information requires all aspects of complete information, and in addition, that all actions taken by participants are known to all others.7 Chess is an example of a game with perfect information if one assumes that chess players are fully capable of remembering all past moves and calculating forward into the indefinite future at any particular stage in the game. Rational players process all available information infallibly. Such players can place past information into long-term storage without loss or bias, and they can bring adequate information into short-term storage to make a correct analysis.

The models of chapters 3–6 assume complete information and infallible processing. In chapters 3 and 4, where there are only two players and each player has only two actions, these assumptions are reasonable, and enable us to make precise predictions. In chapters 5 and 6, where we model eight-player versions of the appropriation game in a finitely repeated setting, the game is more complex but still tractable to analyze. The complexity of the game grows with the number of players, the number of feasible actions, and the number of repetitions (see chap. 3). All this makes the assumption that individuals are perfect and infinite information processors heroic in more complex situations.

Bounded rationality is a much weaker assumption about players' information-processing capabilities. In many situations, the amount of information generated is larger than what individuals can amass and record. They may not utilize all of the information available to them, and they may make errors in processing the information they do use. Boundedly rational individuals possess various heuristics or shortcuts to cope with informationally complex problems they face.

7. Technically, this means that all information sets are singletons and that participants know exactly where they are in a game tree.
Individual Selection Criteria

Theories differ in regard to the criteria they posit actors to follow in making decisions. In many theories that assume complete information and infallible processing, actors maximize expected utility, compute best responses, or obey the minimax criterion. Sometimes these criteria all lead to the same prediction. Usually, however, the predicted outcome is criterion dependent. Under bounded rationality, the information needed for many of the rigorous selection criteria is not assumed to be present. Selection criteria are then built into the heuristics that individuals are posited to use. Sometimes these involve selecting the first alternative that exceeds a minimal threshold. Other heuristics involve more complex processes, but not the necessity of undertaking a full analysis and choosing the maximal set from it.  

Individual Resources

Many theoretical analyses assume that all actors possess sufficient resources to take any of the actions available to them. But in situations where some actions involve high costs, the monetary and time constraints facing individual actors are important constraints. Budgetary constraints may eliminate all but a very narrow band from the feasible set of some actors.

Given that most of the actions that we analyze are feasible to the CPR appropriators, these constraints are not as important in our analysis as they are in many other settings. It does turn out, however, that the amount of the endowment given to subjects in a laboratory setting has a major and unexpected effect on behavior (see chap. 5).

Explaining Behavior in Situations

In order to derive inferences about the likely behavior of each actor in a situation (and, thus, about the pattern of joint results that may be produced), one must make assumptions about the preferences, information-processing skills, selection criteria, and resources of the actors who are participants. The actor is, thus, the animating force that allows the analyst to generate predictions about likely outcomes given the structure of the situation (Popper 1967).

Classical game theory (e.g., Von Neumann and Morgenstern [1944] 1964, chap. 2) assumes that players are fully rational. This has come to mean that players assign complete preferences over outcomes, have unlimited com-

---

8. Tversky and Kahneman 1990 provide a concise overview of research on bounded rationality.
putational powers, conduct complete analyses, and possess the resources necessary for any feasible action. This assumption is intended to apply to both cooperative and noncooperative games. In part because of the extremity of this assumption, powerful mathematical results can be deduced. For many field settings, these theories are highly successful explanatory and diagnostic tools. Even if individuals do not initially behave as predicted, their behavior tends to converge toward predicted behavior over time. For those settings, using these assumptions about individual choice is a useful way of doing institutional analysis. However, empirical results do not always accord with mathematical deductions from these assumptions even after adaptation, learning, or evolution has taken place.

Thus, the theorist has to choose which tools to use to analyze diverse arenas. Within the IAD framework, all of these tools are seen as valuable and having a place in the tool kit of an institutional analyst. The challenge, as we see it, is learning how best to use the full array of tools to undertake theoretical analyses of a wide diversity of situations.

When a theorist analyzes an action arena, specific assumptions are made regarding the structure of the situation and the actors. The task of the theorist is viewed as one of predicting the type of behavior and results, given these assumptions. Questions concerning the presence or absence of retentive, attractive, and/or stable equilibria and evaluations of the efficiency and equity of these results are pursued. The general question being investigated is, given the analytical structure assumed, how does this situation work to produce outcomes?

Evaluating Outcomes

After predicting and explaining outcomes, policy analysts evaluate the outcomes achieved using a diversity of evaluative criteria. The key question addressed in an evaluative effort is: How do predicted outcomes conform to evaluative criteria? As we mention in chapter 1, we rely to a large extent on evaluation criteria related to the concepts of efficiency and Pareto optimality. When individuals craft their own rules, they are apt to rely on additional criteria. Conceptions of fairness are extremely important in deciding upon what type of rules will even be considered as appropriate in a particular community. Whether it is possible for individuals to learn from their mistakes and improve on the outcomes they achieve over time is another important evaluative criterion. Whether rules can be transmitted from one generation to the next without the introduction of substantial error is still another criterion. Thus, there are more criteria to evaluate outcomes than we can rigorously address in this volume.
Factors Affecting Action Arenas

Underlying the way analysts think about action arenas are explicit or implicit assumptions about the rules individuals use to order their relationships, about attributes of a physical world, and about the nature of the community within which the arena occurs (see fig. 2.2). While many analyses are undertaken without an overt attempt to address how these deeper factors affect the situation of interest, theorists interested in institutional questions have to dig deeper to understand how rules combine with a physical and cultural world to generate particular types of situations. Implicit or explicit assumptions about rules, physical variables, and the nature of a community all influence the way the seven elements of an action situation are conceptualized. Thus, an institutional analysis might begin with an analysis of these factors first and proceed to identify some of the typical action situations that result from particular combinations of these factors. Below, we lay the foundation for clarifying the
meaning and attributes of rules, physical conditions, and community. We focus primarily on how rules affect the structure of action situations, since it is rules that are usually the object of efforts to change this structure.

The Meaning of Rules

Rules, as we use the term, are prescriptions that define what actions (or outcomes) are required, prohibited, or permitted, and the sanctions authorized if the rules are not followed (S. Crawford and Ostrom 1993). All rules are the result of implicit or explicit efforts to achieve order and predictability among humans by creating classes of persons (positions) who are then required, permitted, or forbidden to take classes of actions in relation to required, permitted, or forbidden states of the world (E. Ostrom 1986a).

Rules are contextual, prescriptive, and followable (Shimanoff 1980). They are contextual in the sense that they apply to a general set of action arenas but do not apply everywhere. The rules of chess apply only to situations in which participants wish to play chess, and they apply to every such situation. The game of chess provides the context for the application of its rules. Rules are prescriptive in the sense that "those who are knowledgeable of a rule also know that they can be held accountable if they break it" (Shimanoff 1980, 41). Rules provide information about the actions an actor "must" perform (obligation), "must not" perform (prohibition), or "may" perform (permission) if the actor is to avoid the possibility of sanctions being imposed. Rules are followable in the sense that it is possible for actors to perform obligatory, prohibited, or permitted actions as well as it is possible for them not to perform these actions. In other words, it is physically possible for actors to follow or not to follow a rule. This distinguishes actions that are explained by reference to rules or norms from behavior that is explained by the physical characteristics of the situation.

Understanding the relationship between rules and games often requires one to investigate the origin of such rules. In totalitarian governance systems, a central government attempts to impose rules on most action situations occurring within its domain. It attempts to be the source of all rules and their enforcement and invests heavily in police and organized terror mechanisms in this effort. Given the extreme sanctions that can be imposed, individuals interacting with strangers try to stay within the "letter of the law" as prescribed. Behind the scenes, however, many activities are organized using rules other than those prescribed by a central regime (Kaminski 1992). Government officials try to extort bribes from citizens (or businesses), who may try to evade government regulations by keeping some things hidden and paying off officials. Special accommodations are made in secret that are exactly counter to the letter of the law. Thus, in a totalitarian regime where
individuals have had an opportunity to begin to make accommodations with one another, there are many sources of the rules used in daily life. Some of these rules are exactly counter to the prescriptions laid down by the formal government.

In open and democratic governance systems, there are also many sources for the rules that individuals use in everyday life. It is not considered illegal or improper for individuals to organize themselves and craft their own rules, and enforce these rules so long as the activities involved are lawful. Much of the character of law presumes autonomy on the part of diverse, self-organizing patterns of relationships—voluntary associations, families, corporations, municipalities, provinces, and so on. Within private firms and voluntary associations, individuals are authorized to adopt their own specific rules so long as these are within the broad set of potentially lawful rules that are theoretically consistent with the larger constitutional system. Thus, many collective-choice arenas can be used to affect the structure of any particular operational action situation.

When individuals participate in the crafting of multiple layers of rules, some of that crafting will occur using pen and paper. Much will occur, however, as problem-solving individuals interact trying to figure out how to do a better job in the future than they have done in the past. Colleagues in a work team are crafting their own rules when they say to one another something like: “How about if you do A in the future, and I will do B, and before we ever make a decision about X again, we both discuss it and reach a joint decision?” In a democratic society, problem-solving individuals do this all the time. Individuals also participate in less fluid decision-making arrangements, including elections to select legislators. Elected representatives may then engage in open, good-faith attempts to solve a wide diversity of problems brought to them by their constituents. It is also possible in a governance system where individuals are elected, for patterns to emerge that are not strictly problem solving. Incentives exist to create mechanisms whereby one set of individuals dominates over others.

Thus, in undertaking an institutional analysis relevant to a field setting, one needs first to understand the working rules that individuals use. Working rules are the rules used by participants in ongoing action arenas. They are the set of rules to which participants would refer if asked to explain and justify their actions to fellow participants. Rules-in-use may be remembered by participants in local sayings like the Thakali rhyme mentioned in chapter 1. While following a rule may become a “social habit,” it is possible to raise to

---

9. It is not always the case, however, that participants will explain their actions to outsiders the same way they will explain them to fellow participants. Consequently, learning about the working rules used in a particular CPR may be very difficult.
conscious awareness the rules used to order relationships. Individuals can consciously decide to adopt a different rule and change their behavior to conform to such a decision. Over time, behavior in conformance with a new rule may itself become habitual (see Shimanoff 1980; Toulmin 1974; Harré 1974). The capacity of humans to use complex cognitive systems to order their own behavior at a relatively subconscious level makes it difficult for empirical researchers to ascertain what the working rules are for an ongoing action arena.

Rule following or conforming actions are not as predictable as biological or physical behavior explained by physical laws. Rules are formulated in human language. As such, rules share the problems of lack of clarity, misunderstanding, and change that typifies any language-based phenomenon. Words are “symbols that name, and thus, stand for classes of things and relationships” (V. Ostrom 1980, 312). Words are always simplifications of the phenomenon to which they refer (V. Ostrom 1994).

The stability of rule-ordered actions is dependent upon the shared meaning assigned to words used to formulate a set of rules and how they will be enforced. If no shared meaning exists when a rule is formulated, confusion will exist about what actions are required, permitted, or forbidden. Regularities in actions cannot result if those who must repeatedly interpret the meaning of a rule within action situations arrive at multiple interpretations. Because “rules are not self-formulating, self-determining, or self-enforcing” (V. Ostrom 1980, 312), it is human agents who formulate them, who apply them in particular situations, and who attempt to enforce performance consistent with them. Even if shared meaning exists at the time of the acceptance of a rule, transformations in technology, in shared norms, and in circumstances more generally change the events to which rules apply. “Applying language to changing configurations of development increases the ambiguities and threatens the shared criteria of choice with an erosion of their appropriate meaning” (V. Ostrom 1980, 312; see also V. Ostrom 1994, chaps. 1 and 6 for an in-depth development of this thesis).

A myriad of specific rules are used in structuring complex action arenas. Classification of these rules in a theoretically useful typology is a necessary step in developing a cumulative body of knowledge about the effects of rules. Anyone attempting to define a useful typology of rules must be concerned that the classification is more than a method for imposing superficial order onto an extremely large set of seemingly disparate rules. Asking how rules affect the structures of action situations is the method developed as part of the IAD framework to cluster rules. This is seen as a first step in a theory about how rules relate to the structure of action situations, thereby affecting the way individuals behave and achieve outcomes. A similar method can be used in
identifying those aspects of the physical and cultural world that affect behavior and outcomes.

Types of Rules and Rule Configurations

From sets of physically possible actions, outcomes, payoffs, decision functions, information, positions, and participants, rules alter the feasible sets of the values of these variables. The action situation is the intersection of these feasible sets. In regard to driving a car, for example, it is physically possible for a 13 year old to drive a car at 120 miles per hour on a freeway. If one were to model the action situation of a freeway in a state with well-enforced traffic laws, one would posit the position of licensed drivers filled by individuals 16 and over traveling an average of 60 to 65 miles per hour (depending on the enforcement patterns of the state). The values of the variables in the action situation are constrained by the type of physical world involved and then, further affected by the rules-in-use. Most formal analysis of a game focuses primarily on the structure of an action situation: this is the surface structure of our formal representations. The rules are part of the underlying structure that shapes the representations we use. But, how do we overtly examine this part of the underlying structure? What rules should be examined when we conduct analysis at a deeper level?

We identify seven broad types of rules that operate configurally to affect the structure of an action situation. In the list of rules we present here, we emphasize the working part of an action situation (game) that a particular kind of rule directly affects.10

1. Position rules specify a set of positions and how many participants are to hold each position.

   EXAMPLE: Farmers who constitute an irrigation association designate positions such as member, water distributor, guard, member of a tribunal (to adjudicate disputes over water allocation), and other officers of the association.

2. Boundary rules specify how participants enter or leave these positions.

---

10. In this effort, we concentrate primarily on the direct effects of rules (or the physical and cultural factors affecting an action situation). Since all of these factors operate configurally, the final constellation of elements in an action situation depends on more than just one rule per element. The information available to an individual at a node, for example, is directly affected by information rules but also affected by the sequence of activities that are part of an authority rule. One cannot know the action that someone else takes if they must take their action simultaneously with one's own.
EXAMPLE: An irrigation association has rules that specify how a farmer becomes a member of the association and the qualifications that individuals must have to be considered eligible to hold a position as an officer of the association.

3. Authority rules specify which set of actions is assigned to which position at each node of a decision tree. EXAMPLE: If a farmer challenges the actions taken by another farmer or the water distributor, the rules of an irrigation association specify what a water distributor or guard may do next.

4. Aggregation rules specify the transformation function to be used at a particular node, to map actions into intermediate or final outcomes. EXAMPLE: When a decision is made at a meeting of an irrigation association about changing association rules, the votes of each member present and voting are weighted (frequently each vote is given equal weight, but it may be weighted by the amount of land owned or other factors) and added. When 50 percent plus one of those voting (presuming a quorum) vote to alter legislation, the rules are altered. If less than 50 percent plus one vote for the change, the rules remain unchanged.

5. Scope rules specify the set of outcomes that may be affected, including whether outcomes are intermediate or final. EXAMPLE: Rules that specify that the water stored behind a reservoir may not be released for irrigation if the level falls below the level required for navigation or for generating power.

6. Information rules specify the information available to each position at a decision node. EXAMPLE: Rules that specify that the financial records of an irrigation association must be available to the members at the time of the annual meeting.

7. Payoff rules specify how benefits and costs are required, permitted, or forbidden in relation to players, based on the full set of actions taken and outcomes reached. EXAMPLE: Rules that specify whether a farmer may sell any of the water received from an irrigation system, what crops may be grown, how guards are to be paid, and what labor obligations may be involved to keep the system maintained.

The wide diversity of rules that are found in everyday life could be classified in many ways. The IAD method has several advantages. First, rules are tied directly to the variables of an analytical entity familiar to all game theorists. Second, one has a heuristic for identifying the rules affecting the
structure of that situation. Finally, one has a conceptual tool for inquiry about how rules affect a given situation. For each variable identified in the action situation, the theorist interested in rules needs to ask what rules affect the variable as specified. For example, in regard to the number of participants, the analyst asks: Why are there \( N \) participants? How did they enter? Under what conditions can they leave? Are some participants forced into entry because of their residence or occupation?

In regard to the actions that can be taken, the analyst asks: Why these actions rather than others? Are all participants in positions assigned the same action set? Or, is some convener, or other position, assigned an action set containing options not available to other participants? Are sets of actions time or path dependent?

In regard to the outcomes that can be affected, the analyst asks: Why these outcomes rather than others? Are the participants all principals who can affect any state variable they are defined to own? Or, are the participants fiduciaries who are authorized to affect particular state variables within specified ranges but not beyond? Similar questions can be asked about each variable overtly placed in a model of an action situation.

Answers to these sets of questions are formalized as a set of relations that, combined with the structure of a physical world and the type of community involved, produces the particular values of the variables of the situation. As we show in chapter 4, a particular model of a situation could be produced by different underlying factors. Given the frequency of situations with the structure of a Prisoner’s Dilemma, for example, it is obvious that the structure of this action situation results from many different combinations of rules, physical variables, and attributes of a community. This many-one relationship is not problematic when one focuses exclusively on predicting behavior within the one situation. The flip side of this relationship, which is a one-many relationship, is extremely problematic if one were to want to change the situation. From the action situation alone, one cannot infer the underlying factors.

Besides providing a general heuristic for identifying the relevant rules that affect the structure of a situation, a second advantage of examining the rules that directly affect the seven components of an action situation is that doing so leads to a relatively natural classification system for sets of rules. Classifying rules by what they initially affect enables us to identify rules that all directly affect the same working part of the situation. Specific rules used in everyday life are named in a nontheoretical manner—frequently referring to the number of the rule in some written rule book or piece of legislation. Theorists studying rules tend to name the rule they are examining for some feature related to the particular type of situation in which the rule occurs. In
the interests of systematic cumulation, rules structurally the same but called by different names need to be classified the same.

Attributes of a Physical World

The variables of an action situation are also affected by attributes of the relevant physical world. The physical possibility of actions, the producibility of outcomes, the linkages of actions to outcomes, and the knowledge of actors all depend on the physical world and its transformations. The same rule configuration may yield entirely different types of action situations depending upon the types of events in the physical world being acted upon by participants. The difference between goods that are subtractive in nature, such as CPRs and private goods, as contrasted to those that are not subtractive, such as public goods and toll goods, strongly affects how rules affect outcomes. Allocation rules that are essential to achieve better outcomes related to CPRs make no difference in situations where goods are not subtractive. As we discuss in chapter 14, whether a CPR has storage facilities and whether the resource units are mobile also makes a substantial difference in the kinds of rules that one can utilize.

The physical attributes of the relevant world are explicitly examined when the analyst self-consciously asks a series of questions about how the world being acted upon in a situation affects the outcome, action sets, action-outcome linkages, and information sets in that situation. The relative importance of the rule configuration and the physical world in structuring an action situation varies dramatically across different types of action situations. The rule configuration almost totally constitutes some games, like chess, where physical attributes are relatively unimportant. There is little about the size of a chessboard or the shape of the pieces that contributes to the structure of a chess game. On the other hand, imagine, for a moment, switching the balls used in American and European football. The strategies available to players in these two games, and many other sports, are strongly affected by the physical attributes of the balls used, the size of the field, and the type of equipment.

The relative importance of working rules to physical attributes also varies dramatically within action situations considered to be part of the public sector. A legislature is closer in many respects to chess than to football. Rules define and constrain voting behavior inside a legislature more than the physical world. Voting can be accomplished by raising hands, by paper ballots, by calling for the ayes and nays, by passing before an official counter, or by installing computer terminals for each legislator on which votes are registered. In regard to organizing communication within a legislature, however, attributes of the world strongly affect the available options. The physical limit that only one person can be heard and understood at a time in any one forum
strongly affects the capacity of legislators to communicate effectively with one another (see V. Ostrom 1987).

Attributes of a Community

A third set of variables that affect the structure of an action arena relates to the community in which an action situation is located. The attributes of a community that are important in affecting the structure of an action arena include generally accepted norms of behavior, the level of common understanding about action arenas, the extent to which the preferences are homogeneous, and distribution of resources among members. The term *culture* is frequently applied to this bundle of attributes.

If children are taught to extend trust to others so long as the others behave in a trusting manner, adults acquire norms of behavior that enable them to accomplish far more in life in their interactions with other trusting individuals than those who are taught to distrust others in their interactions with other nontrusting individuals.11 This is especially true in relatively homogeneous communities where individuals repeatedly interact with one another along many different dimensions (Taylor 1987). These norms of behavior become a form of social capital that can be drawn on repeatedly as the foundation for cooperative solutions to CPR dilemmas.

Linking Action Arenas

While the concept of a "single" arena may include large numbers of participants and complex chains of action, most of social reality is composed of multiple arenas linked sequentially or simultaneously. Farmers who jointly use an irrigation system, for example, must organize a variety of provision activities primarily related to maintenance. If breaks in the sides of canals are not fixed and the canals themselves not cleaned, the amount of water that actually gets to each farmer's gate declines substantially over time. Organizing the provision side of an irrigation CPR may involve deciding upon how many days a year should be devoted to routine maintenance, how work will be allocated to individual farmers, how emergency repairs should be handled, who is responsible for repairing broken embankments caused by grazing animals, and how new control gates and regulatory devices are to be installed and paid for. Appropriation activities are closely linked to these provision activities. How much water is available for distribution is dependent upon

---

11. It is important to note that the results achieved by individuals who adopted a norm of trusting others depend on the norms adopted by others with whom they come in contact regularly. If the population of "others" contains many who are nontrusting, then those who are trusting may be worse off for following their learned norm.
whether a system is kept in good repair. The level of conflict over water distribution is apt to be higher on a poorly maintained system than on a better-maintained system. In many places in this volume, we will focus on one arena rather than the linked arenas for analytical clarity.

**Multiple Levels of Analysis**

Action arenas are also linked across several levels of analysis. All rules are nested in another set of rules that, if enforced, defines how the first set of rules can be changed. The nesting of rules within rules at several levels is similar to the nesting of computer languages at several levels. What can be done at a higher level will depend on the capabilities and limits of the rules (or the software) at that level and at a deeper level. Changes in the rules used to order action at one level occur within a currently “fixed” set of rules at a deeper level. Changes in deeper-level rules usually are more difficult and more costly to accomplish, thus increasing the stability of mutual expectations among individuals interacting according to a set of rules.

It is useful to distinguish three levels of rules that cumulatively affect the actions taken and outcomes obtained in any setting (Kiser and Ostrom 1982) (see fig. 2.3).

1. *Operational rules* directly affect day-to-day decisions made by the participants in any setting.

2. *Collective-choice rules* affect operational activities and results through their effects in determining who is eligible and the specific rules to be used in changing operational rules.

3. *Constitutional-choice rules* affect operational activities and their effects in determining who is eligible and the rules to be used in crafting the set of collective-choice rules that in turn affect the set of operational rules.

At each level of analysis there may be one or more arenas in which the types of decisions made at that level will occur. The elements of an action situation and of an actor are used to construct these arenas at all three levels. As we discuss above, the concept of an “arena” does not imply a formal setting, but can include such formal settings as legislatures, governmental bureaucracies, and courts. Policy-making regarding the rules that will be used to regulate operational-level action situations is usually carried out in one or more collective-choice arenas as well as being enforced at an operational level. Dilemmas are not limited to an operational level of analysis. They frequently occur at the collective-choice and constitutional levels of analysis.
Multiple Levels of Analysis and Solving Higher-Order Dilemmas

In chapter 1, we discussed two types of coordinated strategies that enable appropriators to extricate themselves from CPR dilemmas. One type of coordinated strategy exists within a set of preexisting rules. The second type of coordinated strategy is an effort to change the rules themselves by moving to a collective-choice or constitutional-choice arena. The possibility of switching to collective-choice or constitutional-choice arenas is frequently ignored in current analyses.

This lack of attention to the possibility of changing the rules of a game results from two different views. The first is a methodological position that eliminates analysis of structural change while examining the effects of one structure on outcomes. In other words, the givens that one uses to specify a problem for analysis are not to be changed in the process of analysis. It is easy to overcome this limit by overtly taking a long-term perspective. It is with a long-term perspective that the given constraints of a particular physical facility are changed into variables that can be changed and thus analyzed. A similar approach can be taken with rules.

The second reason that the possibility of individuals changing their own rules has been ignored is the assumption that a set of rules is itself a public good. Once developed, the rules are available to all individuals, whether or
not they contribute to the effort to design the new rules.\textsuperscript{12} Thus, changing rules is a higher-order, supply-side provision problem (Feeny 1988, 1993). Agreement on better rules affects all individuals in the group whether they participate in the reform effort or not. The temptation to free ride in the effort to craft new rules may be offset by the strong interest that most appropriators have in ensuring that their own interests are taken into account in any set of new rules. Further, the group might be “privileged” in the sense that one or a very small group of individuals might expect such a high return from provision that they pay the full cost themselves (Olson 1965).

The higher the order of a CPR dilemma, the more difficult it is to solve. An even higher-order dilemma than a rule change is the dilemma involving monitoring and sanctioning to enforce a set of rules. Even the best rules are never self-enforcing. It is usually the case that once most appropriators follow these rules, there are strong temptations for some appropriators to break them. If most farmers take a legal amount of water from an irrigation system so the system operates predictably, each farmer will be tempted from time to time to take more than a legal amount of water because his or her crops may be in severe need of more water. Once some farmers take more than their allotment of water, other farmers will be tempted to do the same, and the agreed-upon set of rules can crumble rapidly. Monitoring each other’s activities and imposing sanctions on one another are costly activities. Unless the rewards received by the individual who monitors and sanctions someone else are high enough, each potential monitor faces a situation where not monitoring and not sanctioning may be the individually preferred strategy even though everyone would be better off if that strategy were not chosen. Thus, designing monitoring and sanctioning arrangements that sustain themselves over time is another delicate, difficult, higher-order, supply-side provision task involved in transforming a CPR dilemma.

In many theories of collective action, rules are enforced by outsiders. It is important to inquire into how rules are enforced and how much sanctioning deviant behavior costs. Appropriators may punish one another in several different ways.\textsuperscript{13} In relatively small groups that interact with one another on a wide diversity of fronts, appropriators may impose social sanctions and openly criticize the offender. Further, they may refuse to participate in other types of economic exchanges with an offender. If most members of a community refused to cooperate with an offender, for example, the costs to an

\textsuperscript{12} It is, of course, possible for a group to change their rules and produce a worse outcome. Whether rule changes improve or worsen the outcomes of a situation, the important characteristic of the provision of rules is that everyone using a CPR is affected whether or not they spent time and effort in devising a new rule system.

\textsuperscript{13} See Jankowski 1991 for a discussion of the costs of using different types of punishment mechanisms and how these costs may vary with the size of a group.
offender may be rather substantial, while the costs to other members of the community may be relatively modest. Actions that involve physical coercion or the impoundment of property represent a somewhat higher level of cost for the person who undertakes sanctioning. When several individuals jointly undertake the sanctioning, however, individual costs are reduced and costs imposed on the offender can be very high.

A third type of punishment involves a form of retaliation whereby individuals stop abiding by the rules they have established for some time so as to “teach” the offenders the costs of breaking agreements. This is the type of punishment posited as the means of solving an infinitely repeated social dilemma game. The term used to describe this type of punishment is trigger strategy. What is meant by this term is that appropriators would stand ready to retaliate with strategies leading to suboptimal outcomes at any point that someone consciously or by error adopted such an action themselves (see chap. 3). Punishment by withholding cooperation can be very costly for all involved since those who take action also reap lower payoffs.

Conclusions

Substantial theoretical issues are involved in undertaking institutional analyses of any complex and important set of problems, such as those related to the study of common-pool resources. Without recognizing that theoretical languages are nested from the most specific to the most general, scholars conducting work at one conceptual level may not recognize the array of alternative conceptualizations that are potentially possible and useful for analysis at a particular level. Because one model provides insight into a particular problem does not preclude the possibility of alternative models that usefully illuminate the problem as well, in some cases leading to complementary insights. Having alternative models enables one to carefully specify variable contingencies pertaining to empirical work. Alternative models generate competing hypotheses that can then be tested. Similarly, alternative theories may be needed to address different types of situations that look initially as if they are the same. Because one theory is more useful for some situations than another does not negate the potential usefulness of both theoretical explanations. Similarly, there are many frameworks that can potentially be used in the social sciences, but given the organizing character of a framework it is more difficult for scholars to work across different frameworks.

In this book, we rely on the IAD framework as our general organizing mode that can be used to orient oneself to a large variety of problems. We started with one theory—noncooperative game theory—as our primary tool to construct models of diverse CPR situations. We have found this theory to be an extremely useful and powerful tool throughout our research effort. We
have, however, found empirical evidence that we cannot explain with the initial theoretical tool that we adopted. Thus, we have added a complementary tool—the theory of bounded rationality—to the set of theories we use for explaining behavior related to CPRs. We do not address the question of alternative frameworks within this volume as we have not yet encountered problems where the IAD framework is not a useful tool for addressing policy problems. However, we do not presume that the IAD framework is the only framework available to social scientists interested in understanding questions of social order.