Human choice and climate change

VOLUME THREE

The tools for policy analysis

EDITED BY

Steve Rayner
Elizabeth L. Malone

Pacific Northwest National Laboratory

Battelle Press

CHAPTER 2

Games and simulations

Edward A. Parson & Hugh Ward

In the previous chapter we saw how economic discussions of decisionmaking are deeply rooted in methodological assumptions of atomistic rationality; that is to say, they tend to treat decisions as if they are made by a unitary actor based on rational expectations. Hence, as Chapter 3 describes, the kind of information that is generated to help societal decisionmaking about climate change (or, indeed any other issue of public policy concern) is often structured as if addressed to a benevolent dictator seeking to make the most efficient decision on behalf of us all. In this chapter, we begin to explore the limitations of analysis based on this kind of assumption. We use game theory and simulations to introduce the problem of how multiple decisionmakers acting in their own self-interest may or may not produce an outcome that is rational from a global standpoint.

The first half of the chapter discusses this problem from the standpoint of formal game theoretic analyses. This body of social science research and writing explores decisionmaking among limited numbers of unitary rational actors. In game-theoretic explorations of international relations, these actors are nation states. Although one-shot games are recognized as having very limited application to ongoing relations among states, iterated games do provide a parsimonious framework for thinking about cooperation and decisionmaking at levels between the benevolent dictator, on the one hand, and the anonymous market characterized by many well-behaved individuals, on the other.

However, notwithstanding efforts to develop nested or two-level games (e.g., Putnam 1988), game-theoretic approaches are generally not so well adapted to explore the effect on decisionmaking of nonunitary actors; that is, they do not take account of the tensions among rival viewpoints and values within a state that can cause it to change course during negotiations in ways that cannot be predicted within the parameters of the game. This is one of the issues that is used in the second half of the chapter to justify supplementing formal game theoretic analyses with actual simulations involving human actors representing diverse interests within teams of players representing national positions in iterated cooperation games.

Game theoretic approaches to international cooperation

Global environmental change issues raise the question of international cooperation and collaboration to overcome the problems associated with them. In contrast to local environmental questions that affect specific regions or countries, global environmental change results from activities by individuals, firms, social groups, or entire countries—activities that have global consequences. This is true in particular for climate change, where local emissions of greenhouse gases, resulting from a variety of human activities, have global effects: the mixing of these gases in the atmosphere is so thorough that they may contribute to global climate change by increasing the greenhouse effect on the Earth. This process means that there is no a priori relation between the quantity of greenhouse gases that a region or a country emits and the climate change consequences that it experiences as a result of these emissions.

Hardin's metaphor of the "tragedy of the commons" (1968), in which self-interest and the lack of any constraints on access lead to the overexploitation of open access grazing seems, at first sight, to be a useful way of thinking about the dilemma faced by the international community on global climate change.

Stabilization of the global climate system can be conceptualized as a relatively pure international public good: nations not paying for the cost of stabilizing emissions cannot be excluded from the benefits of a stable climate; climate stability is a good in joint supply, because all countries can enjoy it without prejudice to others' consumption (Weale 1992). Framed thus, the problem is that, because the benefits cannot be limited to those who pay for stabilization, countries may rationally free ride, that is, take advantage of the benefits produced by sacrifices (made by other nations) at no cost to themselves.

National self-interest seems to pressure many nations toward free riding, so that we are currently failing by a wide margin to do what may be required for long-run stability of greenhouse-related emissions. However, just as many small communities over the millennia have developed institutions that have prevented the tragedy of the commons from occurring (Berkes 1989, Ostrom 1990), many hope that the international community will develop the necessary institutions and agreements to restrain the pursuit of national interests.

Game theoretic models offer a way to examine issues of international cooperation, negotiation, and bargaining—especially in the context of international public goods. One of the assumptions included in the practical use of game theoretic models is that participants in international interactions (either nations, or subnational or transnational groups) can be viewed as unitary actors making choices between strategies so as to maximize their expected payoffs. This assumption is made mostly for practical reasons. Elaborated game theoretic models can be constructed from the bottom up, starting with individuals or small groups, and then generate preferences for large groups as well as national preferences. However, because of their size and complexity, models constructed in this way would be excessively difficult to handle.

Generally, a nation's payoff from adopting a particular strategy will vary, depending on the strategies chosen by other nations. To make a rational choice among strategies, a nation has to be able to predict the responses of other nations. The simplest models assume that nations know not only their own payoffs but also those of all the other nations or groups. Also, they assume that all nations are rational and known to be rational. Thus, nations can predict the responses that others will make to any strategy that they choose. An *equilibrium* is a strategy vector where each nation's strategy is a best response to what the others are doing. The prediction is that rational actors will play strategies corresponding to one of the equilibria of the game because, in an equilibrium, no country has an incentive unilaterally to change strategy.

The game theoretic conception outlined above includes the assumption of a priori knowledge of the payoff structure. Clearly, however, in the area of climate change such an assumption is not warranted, since the benefits of greenhouse gas emission restrictions are very difficult to evaluate. The latter occurs, in part, because the effects (damages) associated with global climate change are not yet well known. It has even been suggested that some countries or regions might actually benefit from global climate change (see Oberthür 1993). Therefore, payoffs can only be evaluated in a probabilistic rather than deterministic fashion and conceived of as expected utilities. In principle, resorting to expected utilities to define payoffs and assuming a risk-averse attitude (i.e., emphasizing the dangers and uncertainties of global climate change) should reinforce the precautionary principle and lead actors to cooperate in taking emission reductions. However, the precautionary principle is contested by a school of thought that stresses the importance of uncertainty and the variance associated with the expected outcome, and not just its mean realization (which is implicitly the way the concept of expected utility works).

Including estimated variance as well as averages to evaluate the likelihood of an outcome is part of the conception put forward by Allais (1953) to assess risky situations. In particular, Allais asserts that individuals avoid outcomes associated with large uncertainties, even if they appear more rewarding than outcomes with little or no uncertainty. The risk-averse nature of actors has also been questioned at the individual level by the studies made by Kahneman et al. (1982), who have noticed sudden reversals in risk preferences. It is unclear how group preferences evolve as a result of risky, uncertain, and potentially detrimental outcomes. If there are as many differences among groups as there are among individuals, their perceptions of risk and uncertainty might strongly affect bargaining strategies and thus outcomes of attempted international cooperative arrangements.

In summary, two major cooperative problems emerge at the international level concerning the environment, in general, and climate change, in particular:

- International cooperation is often needed to achieve a collective good and to create a particular institutional framework to keep free riding from occurring. The collective or public good problem to be solved is similar to a Prisoner's Dilemma situation, in which a detrimental equilibrium is obtained in a one-shot situation but where mutually beneficial cooperation can emerge over time as a result of successful threat of retaliation strategies.
- International cooperation often consists of enforcing rules of mutual restriction, such as the reduction of greenhouse-related emissions. This leads to the dilemma of common aversion outlined below and exemplified by the game of Chicken, which contains several equilibria. Paradoxically, such a situation might be more difficult to solve because of the ineffectiveness of retaliation threats (see Ward 1993). The question of international cooperation is complicated further by the fact that the two categories for collaboration outlined above often cannot be separated in the analysis of concrete situations. The creation of an international climate change regime involves both the creation of a public good and the establishment of rules for mutual restriction in order to avoid a mutually detrimental outcome.

Game theory has been used by several authors to theorize about the possibilities of international environmental cooperation (e.g., Taylor & Ward 1982, Livingston 1989, Livingston & von Witzke 1990, Maler 1990, Hoel 1991, Ward 1993, Soroos 1994). However, relatively little has been written with specific and detailed application to global climate change. In the following, we will present a simple iterated model from a game theoretic perspective, *supergame*, apply it to global climate change, relate the model to the debate between realism and institutionalism about the role of institutions, and finally raise some issues of institutional design.

One-shot games are widely recognized to be inadequate models of international cooperation, although they provide metaphors for certain failures of collective actions at the international level (Keohane 1984, Snidal 1986). Even if an international agreement has been signed, the possibilities remain that some countries may overtly break away from it or, more or less covertly, fail to implement it. Thus, nations should be pictured as having repeated opportunities over time to make decisions about whether or not to cooperate. They play so-called *supergames* in which they repeatedly play a one-shot game—with the number of rounds being infinite or uncertain. For clarity of presentation, we assume in formal approaches that cooperation refers to positions favoring greenhouse-related emissions reductions and vice versa.

The basic idea of the model is that the players choose strategies so as to maximize the sum of their own supergame payoffs through time. In calculations of this sum, future payoffs weigh less heavily, that is to say, they are time and cost discounted. A supergame strategy consists of a plan of how to play in each future round, given every pattern of play that would have preceded that round. For a formal statement of the supergame model, see Box 2.1.

The key to cooperative collective action in supergames is the possibility of

Box 2.1 The supergame model

The game matrix in Figure 2.1 may represent row and column's payoffs whether they have Prisoner's Dilemma, Chicken, or Assurance preferences. For Prisoner's Dilemma the ordering of the payoffs is as shown in the diagram. If $x>y$ and $w>z$, the player has Assurance preferences. There are two versions of Assurance, depending on whether $w>y$ or $y>w$. If $y>x>z>w$ and $y'>x'>z'>w'$, the game is Chicken.

The players play an infinite number of rounds of the game, discounting future payoffs. For row from the perspective of round 1, a payoff of P gained in round t is worth dtP, a smaller value of d meaning heavier discounting of future payoffs. Column's discount parameter is d'. Players aim to maximize the discounted sum of their payoffs in each round, taken over the infinite number of rounds. Thus, for example, if both players cooperated in each round, row's Supergame payoff is

$$\lim_{t\to\infty} \sum_{t=1}^{t=t^*} (dx + d^2x + d^3x \ldots + d^{t^*}x) = dx1/(1-d)$$

and column's Supergame payoff is $d'x'/(1-d')$.

making the choice of cooperation conditional on the past cooperation of others (Taylor 1987). If others did not cooperate in the past, this triggers retaliation in the form of refusal to continue to cooperate in the future. Conditionally cooperative strategies of this sort embody threats. If the penalty is large enough, it may pay others conditionally to cooperate. In the context of global climate change, an example of such a strategy might be that the European Union would press ahead with cutting its greenhouse-related emissions so long as the other major industrialized economies do the same; but if they fail to cooperate in this way, the European Union would switch its strategy; that is, it would abandon its plan to make further emissions cuts. It is important that the threat built into conditional strategies is credible; credible threats place restrictions on plausible strategies and equilibria (Fudenberg & Tirole 1991).

To illustrate these conclusions, it is assumed that negotiations are bilateral or that two groups of countries contemplate the merits of mutually beneficial agreements (Fig. 2.1). We call the groups or blocs "rows" and "columns". Each side has two strategies: to cooperate in some measure which it is believed will help stabilize the global climate, or not to cooperate. Suppose for the moment

	C	NC
C	x, x'	z, y'
NC	y, z'	w, w'

where $y>x>w>z$ and $y'>x'>w'>z'$

Figure 2.1 The one shot Prisoner's Dilemma Game payoff matrix.

that both countries favor noncooperation over cooperation regardless of the strategy chosen by the other country and the game is played only once—a one-shot game. The resulting equilibrium for this Prisoner's Dilemma game is where both players choose noncooperation. *Pareto-efficient* outcomes are such that there is no alternative that is better for one side without making the other side worse off. Thus, the outcome is not efficient in this sense. As in Hardin's tragedy of the commons (Hardin 1968), the failure of collective action, which is conceived of as the rational pursuit of individual interests, leads to an inefficient outcome. Although this kind of game theoretic analysis is often applied at the national level, it can also be carried out at the group or political movement level. For instance, Hillman & Ursprung (1992) showed how policy coordination between environmentalist green movements can take the form of a Prisoner's Dilemma game and how this inefficient outcome can sometimes be overcome.

In a so-called Chicken game, row may rationally choose noncooperation if column chooses to cooperate, and cooperation if column chooses noncooperation. Column has the same preference pattern. Pure strategies have two equilibria and, in each of these, one side decides to cooperate and the other side decides not to. Each side has an incentive to commit to noncooperation in order to hijack the other side into cooperation (Schelling 1960). Each side can be expected to be tempted toward brinkmanship, swerving only at the last minute, if at all, away from the strategy of noncooperation. One side may swerve, so that the equilibrium is reached where one side free rides and the other cooperates. However, the danger is that both sides cannot reverse commitments from noncooperation to cooperation, again leading to a failure of collective action. It has been argued that Chicken is an example of a dilemma of common aversion in which the key problem is that of coordinating strategies, so that one of the equilibria—which all sides agree is better than both sides not cooperating—emerges (Stein 1982). Although coordination is crucial, to characterize Chicken and related games with multiple equilibria in this way ignores the potential dangers of commitment tactics and brinkmanship.

Beside the Prisoner's Dilemma and Chicken games discussed above, other one-shot games have also been found helpful in general discussions of international cooperation (Oye 1986). One important alternative to Prisoner's Dilemma and Chicken is Assurance, in both variations of which it is rational to choose cooperation if the other side chooses to cooperate, and to choose not to cooperate if that is also the choice of the other side. In the one-shot game, we say that a player has the following preferences:

- Prisoner's Dilemma if it always prefers noncooperation—no matter what the other side does
- Chicken if it prefers noncooperation if the other side chooses cooperation, and cooperation if the other side chooses noncooperation

- Assurance if it prefers noncooperation when the other side chooses noncooperation, and cooperation when the other side chooses cooperation.

One-shot games in which the two sides have different preference patterns are plausible, too (Taylor 1987). For instance, one side might have Chicken preferences and the other side Prisoner's Dilemma preferences.

The one-shot game underlying the supergame may take various different forms when each player has either Prisoner's Dilemma, Chicken, or Assurance preferences. Nevertheless, perpetual cooperation can typically be sustained only by conditional strategies. (The exception is the case in which both players have Assurance preferences.) Consider a case where players are conditionally cooperating. Suppose one side considers free riding by not cooperating in some round. In some subsequent rounds, the other side would punish it by changing its strategy to noncooperation. Whether it would choose to stick with its original strategy of conditional cooperation in the face of this threat depends on:

- the short-term benefits from free riding

versus

- the long-term costs to itself if cooperation breaks down.

In turn, the long-term costs depend on how much weight is attached to the future payoffs relative to current payoffs, that is, they depend on how heavily future payoffs are discounted. There will be an equilibrium in which everyone conditionally cooperates if three conditions are satisfied:

- Gains from short-term free riding are low.
- Penalties per round from the breakdown of cooperation are high.
- Payoffs in future rounds are not too heavily discounted.

Variation in these factors across issue areas and across time may help explain differences in levels of international cooperation (Lipson 1984, Axelrod & Keohane 1986). For example, it is often suggested that cooperation was easier to achieve in relation to stratospheric ozone depletion than it will be in relation to climate change, because the total economic costs of abatement are much higher in the second case.

The conditionally cooperative equilibrium is never the only one. For instance, if the game being repeated is Prisoner's Dilemma, noncooperation is always an equilibrium; and if the game being repeated is Chicken, the picture is not fundamentally altered, since the two possible patterns in which one side free rides on the other through time are always equilibria. In fact, if any Pareto-efficient outcomes are equilibria, there will generally be an infinity of equilibria.

Suppose that two blocs of countries repeatedly play the Prisoner's Dilemma game shown in Figure 2.1. Then the feasible payoffs for the supergame all lie within the shaded region of Figure 2.2 (Fudenberg & Tirole 1991). The average payoff per round if both blocks always fail to cooperate is w for row and w' for column. These payoffs are the security levels of each side. No matter what

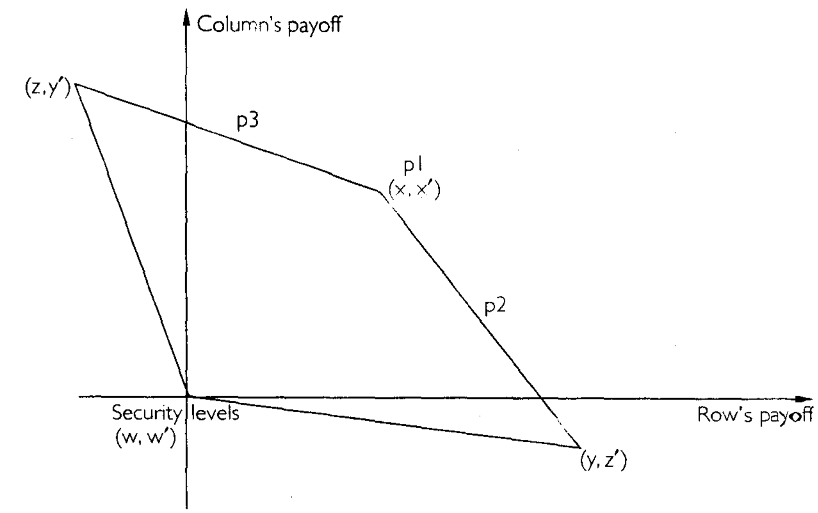

Figure 2.2 Feasible average payoffs

happens, a bloc can never get a lower payoff even if the other side is carrying out a threat against its countries because of their failure to cooperate. The *folk theorem* (so called because no one can recall who first proved it) shows that each payoff point in the shaded region can be equilibrium as long as each bloc puts a high enough weight on future payoffs and each side gets more than its security level (see Fudenberg & Tirole 1991). The intuition is that, so long as sufficient weight is placed on future payoffs to make the punishment substantial and the game lasts long enough, then the threat to drive payoffs down to the security level will deter both sides from breaking away from any pattern of play.

For some commentators, the existence of multiple equilibria calls into question the explanatory power of game theoretic approaches. It may be necessary to resort to an institutional or sociological account of equilibrium selection (see, for instance, Keohane 1988, Sebenius 1992). For other analysts, the existence of multiple equilibria gives explanatory insights into bargaining tactics. The existence of multiple equilibria and conflict of interest over which of those equilibria may combine to produce incentives to use *commitment tactics*. Just as in a one-shot Chicken game, each actor will try to reach an equilibrium with the highest possible payoff. This can be illustrated as follows. Point $p1$ of Figure 2.2 is associated with each side cooperating in every round, getting average payoffs of x and x' for row and column respectively. At $p2$, row cooperates less often. For instance, it might start cooperating after column does, free riding for several rounds on column's actions before it is willing to resume to cooperate. At $p3$, column gets a higher payoff than at $p1$. Row prefers $p2$ to $p1$ to $p3$; column prefers *p*3 to *p*1 to p2. Suppose each of these payoff points can arise in equilibrium. Then row might try to get *p*2 and column might try to get *p*3, each side committing to delaying cooperation until after the other moved. The threats implicit in these strategies of both sides may be triggered, resulting in a worse-all-around outcome in which only security level payoffs of *w* and *w'* are enjoyed. This is analogous to the collision that occurs in one-shot Chicken games when both sides are committed to noncooperation. It has been argued that repeating an underlying Chicken game increases the dangers of noncooperation, because it creates incentives to build and maintain a reputation for toughness (Oye 1986). The same arguments apply to other cases, including the Prisoner's Dilemma game.

Besides the general commitment problem, a general problem of *distrust* also enters in iterated games of qualitatively the same sort as in one-shot Assurance games (Sen 1969, Ward 1989). In order to cooperate, each side has to be assured that the other will also do so. If either believes that there is a large enough probability that the other will not, the first may rationally choose noncooperation rather than risking the worst outcome in which one side cooperates and the other side free rides. Distrust on both sides may be so high that each plays safe by choosing noncooperation. In the supergame, the same problem arises. Assurance may be lacking, because it is suspected that the other side's declarations of intent to cooperate are a tactic to lure the other actor into cooperation with a view to getting a short-term free ride. For instance, the outcome in which both sides always defect must be an equilibrium if the underlying game is Prisoner's Dilemma. Even if players suspect that cooperation in every round is stable, distrust may prevent cooperation from occurring (Ward 1989). The general problem of distrust can arise in other cases, too. Relatively uncooperative equilibria may exist that are worse all round than more cooperative equilibria, and distrust may lead to failure of collective action.

Applying the supergame model to global climate change

Despite its simplicity, the supergame model provides useful insights into global climate change. Many of the conclusions carry over when the model is made more realistic. For instance, there are clearly multiple levels at which nations could cooperate in relation to global climate change, so that the choice is not the binary one of cooperate versus not cooperate. Emission levels could range from further increases, through a freeze, the 20 percent cuts discussed at the Toronto Conference, to the 60 to 80 percent cuts that the IPCC suggests would be required to achieve atmospheric stabilization. In addition, nations might agree to varying degrees of resource transfers to facilitate monitoring and joint implementation, or varying degrees of transfer of control of policy implementation to international agencies. Yet the problems of short-termness, commitment, and

distrust identified in the binary choice supergame model continue so long as the following conditions are all fulfilled:

- There are several discrete levels of cooperation.
- The outcome where all sides cooperate to a high degree is among the efficient outcomes.
- The outcome where all sides cooperate to a high degree is not necessarily an equilibrium.

One interpretation of the current state of play in relation to global climate change politics is that collective action has failed. On paper, signatories to the Framework Convention on Climate Change appear to have moved beyond the cooperative zero point. Moreover, some nations will probably go further, developing policies actually to cut their emissions of greenhouse gases. Even supposing that nations intend to carry out their current commitments, the equilibrium is one where the level of cooperation is generally low, and a case can be made that all-around cooperation at a higher level would be a good collective insurance policy against the risks of global warming. Also, many nations seem to be forgoing national benefits from no-regrets energy efficiency policies. This seems irrational at first sight, yet it may be explained by the desire to gain a reputation for doing little with a view to getting an outcome closer to the national interest in the long term.

The supergame model identifies heavy discounting of future payoffs and uncertainty about benefits as likely causes of the low level of cooperation. Politicians discount future payoffs particularly heavily, because their focus is on the short-run dynamics of support and the reaction of capital markets in which heavy discounting of future investment returns are the norm. The problem of short-termness is exacerbated by a time pattern in which the financial and other benefits from current cooperation arise in the future. Also, uncertainty about the level of future benefits makes risk-averse decisionmakers even less prone to take gambles.

While recognizing the limitations of the FCCC, some see it as a first step to a solution—analogous to the process leading ultimately to the Montreal Protocol. The hope is that levels of cooperation will gradually be increased as scientific certainty and trust between nations increase (e.g., Lang 1993). Distrust is clearly a problem in relation to climate change, just as it was in the case of stratospheric ozone depletion (Ward 1993). From the viewpoint of supergame analysis, a strategy to gradually reduce tension (e.g., Osgood 1979, Ward 1989) may be a way to get from the status quo to a more efficient equilibrium. Nations may be willing to increase their level of cooperation once they see others actually reciprocating cooperation at the current level. Theory suggests that it may pay to make cooperative probes, pushing somewhat further than others to gain valuable information about whether they will reciprocate (Ward 1989). The unilateral policy initiatives to cut greenhouse-related emissions, entered into by some states, may be interpretable in this way, although playing to domestic electoral sentiment and seeking energy efficiency gains are alternative explanations.

Comparing the likely direct abatement costs measured, for simplicity, as the share of gross domestic product (GDP) committed to emission reductions, some nations are currently cooperating more than others. This might result from the unilateral pursuit of no-regrets policies by some and the lack of such easy gains for others. However, the supergame model suggests another explanation: researchers and decisionmakers ought to observe nations committing themselves to relatively low levels of cooperation, both

- to try to bring about a pattern where they currently do relatively little
- to build and to maintain a reputation for tough bargaining.

The actual case of commitment tactics lends some support to this idea (Ward 1993). For instance, while the CANZ group (Canada, Australia, and New Zealand) was willing to take the first steps of setting targets and timetables for stabilizing emissions of greenhouse gases in the late 1980s, the Bush administration in the United States (as part of a block which then included Japan and the former Soviet Union) argued that there was insufficient scientific evidence to justify commitments to emissions reductions, thus ensuring that no specific timetables were written into the FCCC. Even under the more environmentally proactive Clinton administration, the difficulty of steering anything but the most anodyne legislation through Congress effectively binds the United States to relative inaction, even though the administration proposed a plan for stabilization within a definite timeframe. The member states of the European Union may move more rapidly. However, there have been difficulties in the European Union over burden sharing and the carbon tax, partly because of the United Kingdom's refusal to pick up the burden of poorer member states. Again, the Rio Earth Summit saw potentially important actors, such as Brazil, China, and India, committing themselves to inaction unless the North paid a substantial part of the abatement costs. Thus, one incentive to build a reputation for toughness is, over time, to remain part of a larger bloc which can avoid major abatement costs.

The worse the collision outcome resulting from noncooperation (relative to joining in cooperation), the more likely a nation is to back down by switching to cooperation. Nations that stand to lose little from failure, or can make others believe they see things in this way, are in a powerful bargaining position. No matter how much or little climate change affects less industrialized countries, the bargaining power of this group will be enhanced if the strenuous attempts it made in the process of negotiating the FCCC to convince others that it was relatively unconcerned about failure actually work.

Game theory, international regimes of cooperation, and dilemmas of institutional design

Game theory can help to elucidate the various schools of international relations theory and the concept of regimes (Vol. 1, Ch. 6). While some *institutionalists* come close to seeing international law as binding, others have moved closer to the realist assumption that the world is in some sense anarchic (Waltz 1979, Oye 1986, Grieco 1988). However, even if the international system is anarchic, states can cooperate with the assistance of international regimes. Regimes of cooperation consist of formal and informal institutions, shared principles, norms, rules, rights, and decisionmaking procedures. Regimes can provide more favorable circumstances for the existence of conditionally cooperative equilibria, even though they cannot enforce binding agreements. Realists agree that regimes help solve collective action problems, but they are generally more pessimistic about the extent and stability of cooperation (Grieco 1988, Baldwin 1993).

Regimes constrain interdependent decisionmaking by coordinating actions and fostering various forms of collaboration in a way that makes efficient outcomes more likely (Stein 1982). First, regimes may alter the incentives to free ride by threatening to reduce the payoffs for free riders (Axelrod & Keohane 1986, Oye 1986). Second, they provide an institutional context within which a reputation for trustworthy cooperation and for carrying out threats can be built up and then cashed in, both in future rounds and related bargaining forums (Young 1989). Third, monitoring arrangements are typically built into the regime (Levy et al. 1993), and this encourages conditional cooperation by making free riding more visible (Oye 1986, Lipson 1984). Fourth, diplomatic activity on the part of the secretariats of institutions associated with regimes may help to dispel distrust and increase the capacity of nations actually to meet commitments (Levy et al. 1993). Even if a regime has no current value, nations may maintain it because the regime may be useful in the future or because it has attained legitimacy in its own right (Stein 1982, Keohane 1984, Young 1989).

At first sight, a major difference appears to divide realists and institutionalists, because the former emphasize payoff differentials, whereas the latter emphasize absolute payoffs (Powell 1991). Realists argue that institutionalists are too optimistic about the possibilities of cooperation, because they ignore relative gains (Grieco 1988). However, the relative gains perspective opens up difficult issues of interactor comparison of utilities, which might be better treated in the form of a classical but noncooperative gain perspective. Although some have argued that negotiations over environmental problems do not involve relative assessments of payoffs (e.g., List & Rittberger 1991), this seems implausible to realists in the light of the economic and strategic implications of the very large flows of resources involved in moving away from a fossil fuel economy. One argument is that relative payoffs count, because they translate into differentials in future power capacities to remain secure and to alter outcomes (Waltz 1979, Powell 1991). Thus, they affect long-run absolute payoffs. The weight placed on relative payoffs goes up in times of uncertainty and insecurity (Grieco 1988), an argument that may well become pertinent if the fears of some authors about the adverse effects of climate change on international security are realized (Homer-Dixon 1991).

Too much can be made of the apparent differences between the two sides over relative gains. Institutionalists regard regimes as normative orders (Jervis 1988, Weale 1992) in which considerations of fairness have a major impact on states' behavior (Stein 1982, Krasner 1982). This inevitably implies that comparisons of payoffs and relative deprivation matter to nations. Albeit for different reasons, realists and institutionalists game theory suggests that both need to take the relative gains issue seriously.

As relative payoff differences become more and more important, the conditions under which conditionally cooperative equilibria generally exist become more restrictive (Powell 1991, Nicholson 1994; but see also Snidal 1991). As time goes by, an asymmetric equilibrium in which some nations are perceived as cooperating to a much greater extent than others will provoke greater concern for relative gains. It will eventually become apparent that some nations do not honor their obligations. As in the case of burden sharing in the North Atlantic Treaty Organization (Olson & Zeckhauser 1966), there may be growing domestic perceptions of unfairness in nations who shoulder a large part of the collective burden.

This argument suggests that it is important to try to design international regimes in such a way that they steer attention away from asymmetric equilibria and toward equilibria in which no major player gains in relative terms. Bounded rationality (Simon 1982) may make it difficult if not impossible for players to know what the full range of equilibria is or what the best response to others' current strategy is. If institutionalists are right to suggest that international regimes can steer the agenda in relation to problems in the global commons (Keohane et al. 1993, Weale 1992), they may also be able to produce equilibria in which the relative gains perspective is not as much of a problem as a focal point for bargaining (Schelling 1960, Levy et al. 1993, Weale 1992). Although formal theory illuminates the problems here, the question of what ought to be, or might actually be, considered a fair outcome is treated in more detail in Volume 1, Chapter 5.

Despite the prominence of the North–South split in the politics of global climate change, there are arguably more than two bargaining blocs, and there

is evidence that the coalition structure shifted both before and during the Earth Summit (Paterson & Grubb 1992, Nilsson & Pitt 1994, Mintzer & Leonard 1994). A split emerged in the Northern bloc between the United States and other nations, the picture being further complicated by the fact that the United Kingdom, to name one example, often seemed to be close to the United States position. Also, newly developed economies with large fossil fuel reserves and forests, such as China, Brazil, the Organization of Petroleum Exporting Countries (OPEC), and India took a tougher line than others in the South, notably the Association of Small Island States.

However, the conclusions reached from supergame analysis tend to be strengthened if there are more than two blocs of players. First, the commitment problem seems to be even more likely to arise. When the underlying game is a version of Prisoner's Dilemma (Taylor 1987) or a version of Chicken, there is typically a multiplicity of equilibria where some players always free ride and some cooperate in every round. There are additional dangers in attempts by nations to free ride permanently by using commitment tactics when increased numbers make the commitment scramble even more chaotic. Increased numbers also pose difficulties for regimes: problems of distrust are more likely to arise as the numbers of players goes up, because:

- the amount of information necessary to be assured that a nation's cooperation will be reciprocated increases
- it becomes more complex and difficult to apply conditional sanctions (Axelrod & Keohane 1986, Oye 1986)
- transaction costs in deal making rise (Oye 1986)
- the second-order collective action problems surrounding who should punish defectors become harder to solve (Axelrod & Keohane 1986)
- underprovision of compliance mechanisms becomes more likely (Young 1989).

Another reason for pessimism about the chances of collective action in relation to global climate change as compared to the stratospheric ozone depletion problem is the relatively large number of major players in the climate change game. The arguments relating numbers of players to successful collective action make it tempting to go for a less inclusive regime than the one constructed at Rio—or a *fast track* option within the existing convention. Relatively small numbers of like-minded countries (probably members of the OECD, with the United States being a less plausible member of the group) could push cooperation among themselves to relatively high levels (Andresen & Wettestad 1992, Sebenius 1994). One possible assumption underlying this strategy is that once a high level of cooperation is firmly institutionalized, other countries would be pulled in. However, a game theoretic perspective provides no good reason to suppose that the coalition would eventually grow to include all the significant players, as assumed. Given likely equilibria where some nations cooperate and others never cooperate, a point will be reached where it does not pay additional nations to join the group of ambitious emissions reducers. It might be possible to break such a pattern of the nongrowth of the cooperative group by using trade sanctions against those outside the cooperative coalition (Sebenius 1994), but this would require an amendment of the General Agreement on Trade and Tariffs or the World Trade Organization. Once asymmetric cooperation appears permanent, the relative gains effect may arise, leading to the erosion of the cooperative coalition. With more than two players, conditionally cooperative strategies are liable indiscriminately to punish both defectors and cooperators, so that their activation may provoke a general breakdown in cooperation (Oye 1986). The dilemma for institutional design is:

- short-term progress is highly desirable given irreversibilities in the damage being done to the global commons
- such progress may be more likely with less inclusive deals among like-minded countries
- stable cooperation in the long run may require taking the grave risk of holding out for an inclusive deal where all major players are perceived as pulling their weight.

Realists and institutionalists also disagree about the role of leadership in regimes. Although some realists associate leadership with superpower hegemony and see hegemony as a necessary condition for cooperation, for some institutionalists leadership can be provided even in the absence of a hegemonic power in the international system, and leadership is just one factor among others increasing the likelihood of cooperation (Keohane 1984, Snidal 1985, Young 1991, Weale 1992). Both sides accept that leadership is potentially important to the success of regimes. Leaders may provide or distribute selective incentives which go only to other countries which cooperate (Young 1991). Regimes typically produce an array of private goods as well as public goods, and these can be selectively directed to ensure compliance, either by leaders or by regime institutions (Young 1989, Levy et al. 1993). In the context under discussion, these private goods might include technology transfers, payment of monitoring costs, and loans to fund transitions to less polluting technologies. Also, leaders with entrepreneurial skills can put together attractive packages of policies across different issue areas (Young 1991). The idea is that players with different perceptions of the importance of issues can be induced to trade concessions on areas that are of relatively low salience to themselves for a better deal on an issue dimension that they consider important. Commentators on global climate change have already noted dilemmas of institutional design associated with trading concessions across issues. Despite its potential benefits in facilitating progress, such issue trading may result in the agenda becoming impossibly

crowded, leading to transaction cost increases and the sort of long delays observed when the Law of the Sea was being negotiated.

The package deal implicitly proposed at Rio by countries in the South (whereby they concede on global climate change if their demands about the international economic order and development are met) may provoke the emergence of a blocking coalition in the North (Andresen & Wettestad 1992, Sebenius 1994). It may be crucial to success that deals are put together which prevent the emergence of blocking coalitions (Sebenius 1991, Sebenius 1994). Also, power differentials affect the ability of states to get outcomes on the Pareto frontier which asymmetrically favor their interests (Krasner 1991, Sebenius 1992). Ideas about winning and blocking coalitions receive more formal treatment in theories of weighted games (Ordeshook 1986). According to this approach, winning coalitions are inherently unstable when trading concessions across different issue dimensions is possible, because members can be seduced away by a sweeter deal, no matter what the current deal that has been struck (Peterson & Ward 1995). The practical consequences of this are inaction, with negotiations being limited only by nations' rational capacity and information to put together new deals and coalitions.

Game theorists acknowledge that institutional rules and decisionmaking structures may keep issues apart and defuse this problems (e.g., Shepsle & Weingast 1975). The thrust of game theory is, then, further to strengthen the arguments for designing the climate regime, so that it deals sequentially with well-defined issues and encourages package deals only when this seems unlikely to destabilize the whole edifice (Sebenius 1994).

Institutionalists' understanding of regimes places them in a constitutive or mutually constitutive position (Krasner 1982) with respect to the actions of nation states, while realists regard regimes as derivative. For institutionalists, regimes are seen as constraints, facts of life facing nations that may not be dispensed with or ignored, even when there are incentives to do so (Keohane 1988, Young 1989). They are able to alter states' worldview and their preferences (Keohane 1988, Young 1991, Levy et al. 1993). From this perspective, cooperation can literally become a matter of socialization or policy habit, rather than something continually scrutinized for its costs and benefits (Stein 1982, Young 1989). These arguments also make institutionalists more optimistic about the chances of regimes bringing about stable cooperation at relatively high levels over global climate change. They also begin to call into question the utility of formal approaches such as game theory.

The contribution of game theory

Although the supergame model can provide useful insights into international cooperation, in general (Snidal 1986), and global climate change, in particular, its limitations need to be acknowledged. First, supergame analysis has not been extended to cover the cases where players' level of cooperation can vary continuously over several dimensions (Jervis 1988, Sebenius 1992) or current-round payoffs depend on past choices, as they may do in a world where certain forms of environmental damage are irreversible. The ability of nations to rationally pursue national self-interest may be severely limited by distortions in dealing with information and bounds on rational capacity to process it (Jervis 1988). Because of this, some have raised doubts about states' abilities to articulate, communicate, and carry through even simple conditionally cooperative strategies (Lipson 1984, Oye 1986), suggesting the need to further develop models of collective action which assume bounded rationality (Keohane 1984).

Game theory cannot constitute a freestanding explanation, because it takes states' preferences, beliefs, and strategic opportunities as given (Jervis 1988). Existing attempts formally to model how nations' preferences over global climate change arise from domestic political competition (e.g., Ward 1993) are poorly integrated with the structural and systemic basis of states' interests that concern realists (Waltz 1979, Jervis 1988, Lang 1993). In defense of game theory, it can be argued that this approach does not even attempt to explain where preferences, beliefs, and strategic opportunities originate. However, there seems to be no clean break empirically between strategic choice and the processes that mold the underlying parameters of the game. For example, preferences may change during the bargaining process.

In practice, game theorists rarely attempt to model the internal divisions within particular governments, but these sorts of differences are important in negotiating the FCCC (List & Rittberger 1991). This suggests a need to take further and to formalize Putnam's idea (1988) of two-level games, in which national political games are developed within an international game (e.g., Dupont 1994). Presently, internal division is represented more frequently in simulation games where representative national teams can be assembled to include diverse actors (see p. 124 on simulation gaming).

Although arguments about commitment and trust seem empirically relevant to the analysis of negotiations, the supergame model tells us little about the patterns of offer and counteroffer observed in negotiations and the coalitional structures that emerge. One problem is that there are many competing formal models of the bargaining process and associated accounts of coalition formation (e.g., Coddington 1968, Ordeshook 1986), most of which assume quite implausibly that binding agreements can be struck and that the efficient outcomes are

known (Sebenius 1992). Although some progress has been made by using formal bargaining models (Hoel 1991), the most fruitful approaches to bargaining dynamics are likely to be those that are both

- informed by empirical observation and experimental work as well as by game theoretic ideas, and
- do not stick rigidly to standard assumptions, such as perfect information and perfect rational capacity to make decisions (e.g., Raiffa 1982, Sebenius 1992, Sjostedt 1993).

Scholars of international relations have adopted different conceptual lenses to study international cooperation and coordination among countries. In particular, they focus on a variety of levels of analysis and variables of interest articulation, and differ about what information and knowledge base they consider decisive for decisionmaking.

Formal game theoretic analysis shows that efforts at international collaboration will not bring about a unique optimal solution satisfying all countries involved. In particular, as the number of bargaining blocs increases in international negotiations, the more difficult it will become to overcome collective action problems. From a practical standpoint, it may appear to be important to prevent blockage of international negotiations by a single country (under anonymity rules). By analogy to revisions of domestic constitutional law, qualified majority voting might enhance the likelihood of getting effective agreements and avoid weak compromises satisfying the most obstructionist country.

If a country takes the lead in promoting international collaboration to limit greenhouse-related emissions, others may try to resist such moves or even to exploit the position of the leading country. However, a government may rationally adopt such a leadership position for reasons of crucial domestic support and pressures (e.g., from environmental NGOs and business interests) and, in addition, to persuade others to join forces. If the latter case materializes, the joint gains lead to an improved cost–benefit balance for all. As a consequence of organized domestic political pressures, groups of countries (mostly economically well developed and with democratic forms of interest representation) are likely to act as leaders in the international arena of decisionmaking on global environmental accords. Paradoxically, countries having important marginal agricultural sectors and, therefore, being most vulnerable to climate change, are often too weak and too preoccupied with the management of their economies to initiate international cooperation in this area. Given the diversity of factors influencing national positions across countries, we are unlikely to find quick agreements favoring the mobilization of major resources in order to prevent or actively adapt to the likely challenges posed by global climate change. Instead, we should expect the emergence of a cooperative framework lacking precise and costly obligations.

Game theory can provide insight into the bargaining around the FCCC. Game theory cannot stand alone, but it may have a symbiotic relationship with other approaches. It poses important questions about institutional design. Although it does not propose clear-cut solutions to these questions, it can add to the rigor of the debate about these vital issues.

Simulation gaming and its applications

Most formally modeled approaches to human choices cope with multiple decisionmakers in one of two ways: either individual choices are aggregated into market equilibria, subject to the usual assumptions as suggested in Chapter 1; or policy choices are assumed to be made by a unitary national decisionmaker as described already in this chapter. But real policy choices are made neither by absolute and unitary national authorities nor by actors so numerous and well behaved that their collective decisions can appropriately be modeled as markets. Rather, outcomes are often shaped by a combination of institutional, political, strategic, and negotiation processes that involve a few major actors or many. These actors' values and preferences may be unclear or contested, particularly as regards outcomes that are multiattribute, risky, or distant in time. Over matters that require joint decisions, parties' interests may be partly common and partly conflicting. Finally, the range of parties' feasible choices may be ambiguous, poorly known, or changing.

When some of these conditions hold, the most basic uncertainties about a problem may concern behavior, values, preferences, or strategic interaction of choices. In assessing such uncertainties, simulation gaming methods can serve as supplements or alternatives to formal analysis of rational actors in negotiations or markets. Like formal analyses, these methods can help shed light on problems that are both too novel to be dealt with by simple reasoning from precedent or analogy and too high-stakes to be dealt with by ad hoc decisionmaking. Unlike formal economic models or game theoretic analysis, these methods introduce decisionmaking by actual people into the assessment, through the use of human participants who deliberate, negotiate, and act within a simulated decision context that represents the issue or policy problem to be investigated. In other words, the moves in the game are actually decided by various players representing distinct decisionmaking entities and their bounded rationalities, rather than by a single gamemaster or analytic entity externally constructing the rational behavior of each party to a negotiation. The objectives of simulation methods are those of integrated assessment (see Ch. 5): to assemble and interpret knowledge from varied domains to serve the needs of policy

and decision. They have been exercised extensively in other fields; their application to environmental problems is relatively new, but growing.

Problems can most usefully be investigated with simulation gaming when an intermediate amount of knowledge is available—enough to represent its basic structure or building blocks (actors, their interests, authority, relationships, and knowledge, plus relevant knowledge about the world), but not so much that conventional, cheaper, or simpler methods can effectively inform policymaking, making the expense and difficulty of simulations superfluous. Examples of problems well suited to simulation gaming include those in which many complex organizational routines must work together smoothly, particularly under conditions of crisis such as arise in military operations and emergency response (Bracken 1990); and major institutional innovations affecting many actors, such as reform of healthcare systems or developing new institutions for managing global environmental problems.

Simulation gaming methods share two basic characteristics. First, they are all representations of a complex system by a simpler one with relevant behavioral similarity (Brewer & Shubik 1979). The behavioral similarity permits learning about the complex system by manipulating the simpler one. Second, they use participants' reasoning, decisionmaking, and action to represent the reasoning, decisionmaking, and negotiation of the individuals, organizations, or governments whose actions are at issue.

Other aspects of simulation gaming methods can vary widely. Participants' choices must be defined, supported, and constrained in a way that fills in the representation of the policy issue in question. Ways of accomplishing this can include the following:

- textual scenarios that describe the decision setting, focus attention on essential elements, and provide a sense of realism and importance for the proceedings
- an expert control team, whose judgments determine the consequences of participants' decisions
- formal models, which can supplement or replace a control team in this job
- specific analytic tools or general information resources to support participants' deliberations, planning, and negotiations.

Applicability of simulation gaming to global environmental change

Problems of global environmental management exhibit precisely circumstances under which formal economic modeling and game theoretic analysis can provide essential but only partial insights. Important decisions are made interactively in international negotiations. There is substantial disagreement and uncertainty regarding the values at stake, as well as the magnitude, locus, and character of potential threats. Decisions have long consequences. And the most important uncertainties may not based in biophysical knowledge, but rather concern values, behavior, and strategy.

Given these characteristics, simulation gaming assessment methods offer several potential advantages over formal approaches alone:

- They support integration of knowledge from a broader set of domains, including disciplinary knowledge and knowledge embedded in formal models, as well as intuition or judgment of multiple experts
- They investigate questions that are primarily behavioral or strategic, such as processes for developing and maintaining cooperation, coalition formation, and the robustness of strategies to uncertainty or surprise.
- They investigate problems for which preferences and values are contested or obscure (and not easily reconcilable through formal devices such as game theory or multiattribute utility theory).
- They promote more effective communication between assessment and policymaking.

Simulation gaming exercises that incorporate formal models can help improve models' utility and relevance by bringing them into a demanding policy-relevant setting for focused scrutiny and use by their intended audiences.

Major approaches to simulation gaming

Simulations come in various forms, and have been applied to a variety of policy domains. Whereas the term "simulation" refers to any representation, including formal models, simulation gaming always involves human participants, possibly augmented by formal models. Simulations that combine human participants with formal models (*person–machine simulations*) can permit more sophisticated representation of technical, scientific, or economic aspects of a problem, and hence may have advantages in investigating environmental issues. These can also be used to educate participants in the modeled phenomena, or to critique and test the plausibility and usefulness of formal approaches. *All-person simulations* emphasize the decision, negotiation, communication, and information-processing aspects of the problem. These are less able to represent the behavior of complex systems, but offer greater play for judgment and discussion on the feasibility, utility, and potential consequences of particular combinations of decisions.

A few basic dimensions capture much of the relevant variation of design among simulation gaming exercises:

- choice of boundaries: how much of the real problem is included in the simulation, and how much of that is represented by the participants and how much by formal models or rules.
- how tightly participants are constrained, trading off the benefits of a sharply focused simulation addressing a very specific question when participants have limited freedom against the benefits of encouraging creative solutions and insights when participants are allowed wide latitude to improvise or challenge the games' presumptions.
- intensity of time pressure and role identification, which can both be varied to promote a character ranging from calm, detached reflection to intense, engaged crisis decisionmaking.
- representation of simulated time, which can be fast or slow and can either traverse a single path through simulated time or multiple paths under different presumed conditions (Toth 1988a)
- treatment of uncertainty: whether participants see explicit uncertainty when considering their choices; whether the development of events in the simulation is stochastic (or is asserted to be); and whether the equivalent decision situations are faced more than once under different realizations of uncertain events.
- the expertise and seniority of participants, and how closely their simulation roles correspond to their real-life responsibilities.

War games and political–military exercises

The earliest simulations used to inform complex decisions were war games, either played on real battlefields or through representations on scaled playing surfaces. Postwar simulations moved beyond strictly military uses to include related diplomatic and political issues, in pursuit of a synoptic-scale representation of foreign-policy crises. The earliest such diplomatic simulations used teams of players to represent national governments, with a referee team ruling on the plausibility of proposed moves. Later developments included teams representing particular national and subnational organizations, and allowed both diplomatic interactions and the playing out of military confrontations. Similar free-form scenario-based simulations modeling foreign-policy crises are still in use (Goldhamer & Speier 1959, Paxson 1963, Mandel 1985, Kahan et al. 1985, Allen 1987).

The structures of all such simulations are similar. A text *scenario* describes the history and context, defines participants' objectives and resources, establishes the essential focus of the simulation, and provides a vivid, engaging setting in which participants can immerse themselves. Players, in two (usually) or more teams, make decisions on behalf of the organizations, groups, or nations they represent, over a period of several days. Simulated time normally stops while teams deliberate over their decisions, and jumps discretely between decisions. The control team, which normally includes the exercise designers and scenario writers, manages the simulation and determines the consequences of teams' decisions—in effect representing the underlying causal structure of the simulation, plus other actors not represented by teams, and nature. In contrast to earlier war-game traditions, the game's assumptions and the control team's decisions are normally open to challenge and discussion (Jones 1985), particularly during the intensive debriefing that follows the play. The control team's ability to maintain a plausible, vivid, evolving history that responds to participants' choices is crucial to these exercises. As such exercises have developed, the control team has increasingly been assisted by formal models (Bracken 1984).

The primary function of such simulations is to stress and test a complex system of organizations, technology, and routines, a system that cannot anticipate every kind of challenge that it may be called upon to face. Simulations permit critical, reflective examination of system response to hypothetical challenges, including crises that would happen too fast to allow such examination, or even considered responses, while they were happening.

Simulations of politics and international relations

From 1956 to 1972, a project at Northwestern University conducted research on person–machine simulations of politics and international relations (Guetzkow & Valadez 1981). These simulations, although inspired by the practical and operational political–military exercises, sought to develop and test theory, coupling international security relations with domestic political processes. Over time, the project's simulations came to include more actors and more issues, and to make increasing use of formal models—including some large all-machine simulations.

The project's first major simulation was the InterNation Simulation (INS), a simulated five-nation, two-bloc world in which national teams' decisions determined national welfare and security, and political change, through a set of programmed rules that took the place of a control team (Guetzkow 1968, 1981, Modelski 1970, Smoker 1972, Bloomfield 1984, Alker 1985). Later project simulations moved in two directions: toward richer and more detailed person–machine formats, and toward pure computer simulations without human participants (Valadez 1981). Principal among the former was the International Process Simulator (IPS), with roles for national governments, national and multinational industry, international NGOs, and international governmental

organizations (Smoker 1981). Programmed (admittedly rudimentary) rules simulated trade and investment, research and development, and popular and labor unrest, whereas domestic politics was represented by teams playing opposing elites. Subsequently, the GLOBUS project sought to combine IPS's political and economic modeling with systems-dynamic models of resource, energy, and pollution constraints (Bremer 1987).

Adaptive environmental assessment and management

The simulation approaches summarized above use formal models to represent the physical world and hence define and constrain the participants' decision environment. Adaptive environmental assessment and management (AEAM) reverses this relationship between participants and models. AEAM is a process for bringing dispersed expertise to bear on complex and contentious decision problems, particularly concerned with natural-resource management. Participants, including both substantive experts and stakeholders, collaborate to construct a simulation model of the system under dispute, with staff support from a group of modelers and facilitators. The process serves two potential purposes: integrating dispersed knowledge and making it accessible to decisionmakers; and encouraging participants to enter a constructive dialog by focusing their attention on the demanding (and interesting) job of building a representation of the system under dispute. In many cases, the value of the exercise comes entirely from the focused discussion among protagonists that takes place around the modeling exercise, rather than from the model itself (Holling 1978, Sonntag 1986).

Policy exercises

A new class of simulation methods to study complex policy problems, including environment and development issues, has been developed, initially at the International Institute for Applied Systems Analysis (IIASA), under the name *policy exercises*. These share with the political–military exercises the emphasis on prior scenario development, and the structure of playing teams and one control team, but differ in the following respects:

- Policy exercises do not normally represent situations of strong intergroup conflict; in fact, some applications are to problems with no salient conflict, and hence use only one team, or parallel teams engaged in identical or complementary tasks.
- Teams' jobs are normally to create and analyze future or counterfactual

histories, with less active control-team intervention than is typical of political–military exercises.

- The exercises typically cover timescales of years to decades, rather than short-term crisis response.
- They represent time flexibly; rather than simply stepping forward in simulated time, they may work forward to or backward from a specified future endpoint, or go over the same period repeatedly under different conditions.

Policy exercises seek to integrate isolated pieces of knowledge, to build communication bridges between the academic and policy communities, and to encourage longer term, less constrained, and more creative thinking (Brewer 1986, Sonntag 1986, Toth 1988a,b, 1994). Obtaining useful results, as in political–military exercises, normally depends on the use of senior, expert policy participants. Policy exercises have been used to study comprehensive synopses of future world histories (Svedin & Aniansson 1987); to develop scenarios for regional impacts of global climate change (e.g., Jäger et al. 1991); and to study regional and forest-industry response to European acid rain (Duinker et al. 1993).

Stockholm Environment Institute (SEI) Greenhouse Policy Exercise

Following a suggestion made at the 1985 Villach Conference (WMO 1986), the Stockholm Environment Institute sponsored a global climate-change policy exercise in September 1990 at Bad Bleiberg, Austria (Jäger et al. 1991). The 25 participants were experts in climate and environmental policy, although none held a senior policy position at the time of the simulation.

The exercise focused on scenarios of the world in the year 2050. It used no computer models, and no gaming through which interactions of participant decisions determined outcomes. Rather, each of three teams worked in parallel to develop future histories leading from the present to an endpoint scenario that specified either large or small anthropogenic climate change in 2050. Teams did this job twice. The first time, each team's scenario was fixed in advance: while each one stated that sustainable development had been achieved by 2050 (defined by level and distribution of world income, rising life expectancy, and reversal of net deforestation and desertification), each specified a different combination of high or low climate change in Europe and in Southeast Asia. In the second round, each team chose its own 2050 endpoint, from a menu that combined different levels of climate change and of world equity, but did not specify particular regional results.

The results of this exercise served primarily to illuminate the participants' assumptions and heuristics about climate change, and the confidence with which these were held. In particular, all teams equated high or low climate

change in 2050 with high or low anthropogenic emissions, neglecting other sources of uncertainty. In particular, every group with a low-climate-change future assumed that this came about through strong emission limits, and so described how major new international commitments and institutions were developed beginning in the 1990s. Every high-climate-change group described much more delayed changes in policies and institutions, which were eventually driven by grassroots political and religious movements reflecting public response to major climatic shocks of the next century.

RIVM simulation gaming and strategic planning exercises

A project of the National Institute for Public Health and the Environment (RIVM) of the Netherlands is developing simulation gaming exercises for global environment and climate change that draw on RIVM's formal integrated assessment models.

A major exercise under development, the Global Environmental Strategic Planning Exercise (GESPE) (de Vries et al. 1993) includes national teams making sequential decisions setting national investment, energy prices, energy conservation and alternative supply options, recycling, and reforestation. Teams' decisions are taken to represent a combination of government policy and aggregate social response. The structure of participants' decisionmaking will be constrained according to a model of decision processes developed by Mintzberg et al. (1976), while consequences will be determined by a set of dynamic computer models. Preliminary development of GESPE has used the systems dynamics model World 4.0, and it is planned to integrate this exercise with one of RIVM's formal integrated assessment models, either IMAGE 2.0 or TARGETS (see Ch. 5).

Preliminary exercises have been conducted using the simpler model SusClime (de Vries 1995), which presents a two-country world undergoing demographic and energy transitions. Nations manage these transitions by trading energy, by exchanging loans or aid, and, most importantly, by allocating current production among five investment alternatives: goods-producing capital; energy capital for fossil, renewable, and conservation; and population capital, which provides for current consumption and welfare, and at sufficiently high levels brings down population growth rates. New energy sources exhibit learning, while accumulating atmospheric carbon dioxide eventually degrades capital productivity. More tightly or more loosely constrained scenarios can be defined by varying the size of fossil resources, and the intensity of climate change impacts.

The primary objective of this project is to convey information embedded in the models—about dynamics, uncertainties, and interdependencies of global climate change, and related demographic and energy transitions—to policy participants. Work with SusClime has also explored the dynamic behavior associated with players adopting various simple heuristic strategies for their allocation choices.

IIASA climate change policy exercise

A series of policy exercises developed at IIASA seek to combine the advantages of scenarios and formal models in simulated decision environments. A test run of the first exercise involved 25 participants in four national teams, negotiating decisions over national emissions commitments plus associated measures for implementation, in scenarios set in 2005 and 2020. Each team had access to the reduced-form integrated assessment model MiniCAM, through a simple spreadsheet-based graphical interface, to design and execute custom model runs to examine the long-term consequences of their choices.

Test-run behavior and results highlighted the negotiation of obligations, rather than using model projections to evaluate long-run strategies, or design implementation details. The exercise generated sharply focused, cogent arguments illuminating current policy disputes over the form of national targets, the use of punitive sanctions, and joint implementation; generated a few policy proposals of moderate novelty; and, most centrally, provided methodological insights into opportunities and pitfalls of simulation design for global environmental change. Subsequent exercises will focus more narrowly on the relationship between emission obligations and financial transfers, and will use an episodic structure in which participants are sometimes organized into partisan national negotiating teams, and sometimes into neutral expert teams (Parson 1996a).

Uses of simulation gaming in integrated assessment

A conventional dichotomy divides potential uses of simulation methods into research and teaching. Some, although not all, potential uses of simulation gaming for integrated assessment can be categorized as research and teaching. In both domains, the potential value of simulation gaming is clear, but carries significant limitations and risks. Using simulation methods to test general research hypotheses about the behavior of people, organizations, or governments poses serious validation problems, because the simulation's contextual richness normally implies that they involve too many potentially confounding variables and admit insufficient replication (Parson 1996b). But strikingly similar phenomena that sometimes arise in diverse simulations, settings, and participant groups suggest plausibly generalizable patterns. Schelling, for example, reports that it is very difficult to keep a crisis at the boil in any simulation and offers a plausible (and testable) general explanation: that team members differ in how belligerently they wish to act, and that teams accommodate these differences by ordering their possible actions, doing the softest ones immediately and postponing the hardest ones to be contingent on other teams' actions. Consequently, teams normally see only the softest end of their counterparts' range of proposed actions; each thinks the others less belligerent than they actually are, and hence never reveals its own belligerence (Schelling 1987). Such generalizable behavior patterns can arise in simulations, although simulation games alone cannot validate them.

Simulation games can also be powerful vehicles for conveying insights to policymakers, although such use presumes that simulation designers or modelers have indeed discerned deep, important truths that should guide policy decisions. This is a bold presumption, and great caution is warranted before assuring policymakers that there are good grounds for confidence that the insights to be conveyed are true, important, enduring, and not liable to misinterpretation.

Other potential contributions of simulation gaming to complex, ill-posed, and imperfectly understood policy problems do not fit cleanly into the categories of research or teaching. Merely by providing a structured vehicle for critical reflection and dialogue, a simulation exercise can change the way participants think about the problem: help to clarify their views of its scope or importance; shake up preconceptions; or identify plausible and overlooked risks, innovations, contingencies, consequences, or responses (Parson 1996c). Simulation exercises can exhibit new, collective-level outcomes that no individual would have thought of alone, bringing the benefit of a list of things researchers might never have thought of (Schelling 1964). By drawing out concrete implications of key uncertainties, simulation exercises can provide a more thorough identification and more practical ranking of policy-relevant uncertainties than formal analyses alone.

Simulation games, particularly when formal models are used within the simulation, can also support focused communication between analysts and decisionmakers. Immediate benefits of this interaction may take the form of educating the policy community in insights embedded in models, or educating modelers to the needs and priorities of policymakers, hence helping to test or improve the policy relevance and legitimacy of formal models. General improvement of such communication may also bring substantial value of a more diffuse and long-term nature.

Significant risks accompany these benefits, principally bias and overgeneralization from small samples (Levine 1964). Subtle bias in simulation design is an ever-present risk that can push results in particular, predictable directions and diminish the heuristic value of the exercise. Both participants and designers are at risk of overgeneralizing, making too much of a sample of one experience, particularly because simulation experiences can be vivid and compelling. However, this risk may be no more serious than the widespread tendency to do the same thing from small samples of vivid real-life experiences. Both kinds of risks can be greatly mitigated by the use of simulation debriefing, in which participants and designers can question the simulation's relevance and generalizability, present other possibilities, and reflect critically on their experience. These risks are also unlikely to be severe when simulations are used to bound issues, clarify relevant preferences or values, or identify overlooked issues.

In summary, simulation gaming methods have potential value as devices for integrated assessment, as supplements to conventional forms of analysis or sober critical reflection, but not as replacements for them. Simulation methods are immature and developing. Their most evident contributions are as scoping devices, to clarify preferences and values, to promote creative thinking, and to identify overlooked issues. Simulation methods merit further exploration and experimentation, to elaborate and define their potential contributions, to make their contributions more reliable and to permit critical professional evaluation of the contribution of particular simulation exercises and approaches. In pursuing this development it will be important to avoid excessive claims and vain hope of certainties, predictions, or strong verifications that simulation methods cannot hope to achieve.

Conclusion

Thus, game theoretic and simulation gaming approaches both move analysis beyond the atomistic rationality of formal models of individual and market behavior, to encompass small numbers of multiple actors with differing perceptions and incentives. Of the two approaches, formal game theoretic analysis remains closer to the paradigm of universal rationality, and accordingly permits the analyst to predict moves based on his or her analysis of the actor's perception of self-interest. In the simulation gaming approaches, players actually represent social entities larger than their individual selves; this dimension injects a larger human element of complex behaviors into the execution of the game. However, although both approaches move beyond atomistic rationality, they continue to rely on two assumptions:

- that parties rationally perceive and act upon self-interest
- that, because rationality is universal, the decisions or actions of a game analyst or of a player in a simulation are reasonably representative of another player faced with the same information and incentives.

We will explore the basis of both of these assumptions in Chapter 3.

References

Alker Jr, H. R. 1985. Global modeling alternatives: the first twenty years. In *Theories, models, and simulations in international relations: essays in honor of Harold Guetzkow*, M. D. Ward (ed.). Boulder, Colorado: Westview.

Allais, M. 1953. Le comportement de l'homme rational devant de risque: critique des postulats et axiomes de l'école américaine. *Econometrica* **21**, 503–546.

Allen, T. B. 1987. *War games*. New York: McGraw-Hill.

Andresen, S. & J. Wettestad 1992. International resource cooperation and the greenhouse problem. *Global Environmental Change* **2**(4), 277–91.

Axelrod, R. & R. O. Keohane 1986. Achieving cooperation under anarchy: strategies and institutions. See K. Oye (1986).

Baldwin, D. A. 1993. *Neorealism and neoliberalism: the contemporary debate*. New York: Columbia University Press.

Berkes, F. (ed.) 1989. *Common property resources: ecology and community-based sustainable development*. London: Pinter (Belhaven).

Bloomfield, L. 1984. Reflections on gaming. *Orbis* **27**(4), 783–90.

Bracken, P. 1984. Deterrence, gaming, and game theory. *Orbis* **27**(4), 790–802.

—— 1990. Gaming in hierarchical defense organizations. In *Avoiding the brink: theory and practice in crisis management* A. C. Goldberg, D. Van Opstal, J. H. Barkley (eds). London: Brassey.

Bremer, S. A. (ed.) 1987. *The GLOBUS model: computer simulation of worldwide political and economic developments*. Frankfurt: Campus.

Brewer, G. D. 1986. Methods for synthesis: policy exercises. In *Sustainable development of the biosphere*, W. C. Clark & R. E. Munn (eds). Cambridge: Cambridge University Press.

Brewer, G. D. & M. Shubik 1979. *The war game: a critique of military problem solving*. Cambridge, Massachusetts: Harvard University Press.

Coddington, A. 1968. *Theories of the bargaining process*. London: George Allen & Unwin.

Duinker, P. N., S. Nilsson, F. L. Toth 1993. *Testing the "Policy Exercise" in studies of Europe's forest sector: methodological reflections on a bittersweet experience*. Working Paper WP-93-23, Laxenburg, Austria: International Institute for Applied Systems Analysis (IIASA).

Dupont, C. 1994. Domestic politics and international negotiations – a sequential bargaining model. In *Game theory and international relations: preferences, information, and empirical evidence*, P. Allan & C. Schmidt (eds). Cheltenham, England: Edward Elgar.

Fudenberg, D. & J. Tirole 1991. *Game theory*. Cambridge, Massachusetts: MIT Press.

Goldhamer, H. & H. Speier 1959. Some observations on political gaming. *World Politics* **12**(October), 71–83.

Grieco, J. M. 1988. Anarchy and the limits of cooperation: a realist critique of the newest liberal institutionalism. *International Organization* **42**, 51–8.

Guetzkow, H. 1968. Some correspondence between simulations and realities in international relations. In *New approaches to international relations*, M. Kaplan (ed.). New York: St Martin's Press.

—— 1981. International relations theory: contributions of simulated international processes. See Guetzkow & Valadez (1981).

Guetzkow, H. & J. J. Valadez (eds) 1981. *Simulated international processes: theories and research in global modelling*. Beverly Hills, California: Sage.

Hardin, G. 1968. The tragedy of the commons. *Science* **162**, 1243–8.

Hillman, A. L. & H. W. Ursprung 1992. The influence of environmental concerns on the

political determination of trade policy. In *The greening of world trade issues*, K. Anderson & R. Blackhurst (eds). London: Harvester Wheatsheaf.
Hoel, M. 1991. Global environmental problems: the effects of unilateral action taken by one country. *Journal of Environmental Economics and Management* **20**, 55–70.
Holling, C. S. (ed.) 1978. *Adaptive environmental assessment and management*. Chichester, England: John Wiley.
Homer-Dixon, T. F. 1991. On the threshold: environmental change and violent conflict. *Scientific American* **16**, 76–116.

Jäger, J., N. Sonntag, D. Bernard, W. Kurz 1991. *The challenge of sustainable development in a greenhouse world: some visions of the future* [report of a policy exercise at Bad Bleiberg, Austria, September 2–7, 1990]. Stockholm: Stockholm Environment Institute.
Jervis, R. 1988. Realism game theory and cooperation. *World Politics* **40**, 317–49.

Kahan, J. P., W. M. Jones, R. E. Darilek 1985. *A design for war prevention games*. Rand Note N-2285-RC, Rand Corporation, Santa Monica, California.
Kahneman, D., P. Slovic, A. Tversky 1982. *Judgment under uncertainty: heuristics and biases*. Cambridge: Cambridge University Press.
Keohane, R. O. 1984. *After hegemony: cooperation and discord in the world political economy*. Princeton, New Jersey: Princeton University Press.
—— 1988. International institutions: two approaches. *International Studies Quarterly* **32**, 379–96. [Reprinted in *International institutions and state power*, R. O. Keohane (ed.). Boulder, Colorado: Westview, 1989.]
Keohane, R. O., P. M. Haas, M. A. Levy (eds) 1993. The effectiveness of international institutions. In *Institutions for the Earth: sources of effective international environmental protection*, P. M. Haas, R. O. Keohane, M. A. Levy (eds). London: MIT Press.
Krasner, S. D. 1982. Structural causes and regime consequences: regimes as intervening variables. *International Organization* **36**, 185–205.
—— (ed.) 1983. *International regimes*. Ithaca, New York: Cornell University Press.
—— 1991. Global communications and national power: life on the Pareto frontier. *World Politics* **43**(3), 336–66.

Lang, W. 1993. International environmental cooperation. In *International environmental negotiations: process issues and contexts*, S. Sjostedt et al. (eds). Stockholm: Swedish Council for Planning and Coordination of Research and the Stockholm Institute of International Affairs.
Levine, R. A. 1964. Crisis games for adults. Internal Research Memorandum, Rand Corporation, circulated summer 1964 and published in *Crisis games 27 years later: plus c'est déja vu?* P-7719. Santa Monica, California: Rand Corporation.
Levy, M. A., R. O. Keohane, P. M. Haas 1993. *Improving the effectiveness of international environmental institutions*. In *Institutions for the Earth: sources of effective international environmental protection*, P. M. Haas, R. O. Keohane, M. A. Levy (eds). Cambridge, Massachusetts: MIT Press.
Lipson, C. 1984. International cooperation in security and economic affairs. *World Politics* **37**, 1–23.
List, M. & V. Rittberger 1991. Regime theory and international environment management. In *The international politics of the environment*, A. Hurrell & B. Kingsbury (eds). Oxford: Oxford University Press.
Livingston, M. 1989. Transboundary environmental degradation: market failure, power and instrumental justice. *Journal of Economic Issues* **23**, 79–91.
Livingston, M. & H. von Witzke 1990. Institutional choice in transboundary pollution. *Society and Natural Resources* **3**, 159–71.

Maler, K. 1990. International environmental problems. *Oxford Review of Economic Policy* **6**, 80–108.
Mandel, R. 1985. Professional-level war gaming: a critical assessment. In *Theories, models and*

simulations in international relations: essays in honor of Harold Guetzkow, M. D. Ward (ed.). Boulder, Colorado: Westview.
Mintzberg, H., D. Raisinghani, A. Theoret 1976. The structure of unstructured decision processes. *Administrative Science Quarterly* **21**, 246–75.
Mintzer, I. & J. A. Leonard 1994. *Negotiating climate change: the inside story of the Rio Convention*. Cambridge, Massachusetts: Cambridge University Press.
Modelski, G. 1970. Simulation, realities and international relations theory. *Simulation and Games* **1**(June), 111–34.

Nicholson, M. 1994. Interdependent utility functions: implications for cooperation and conflict. In *Game theory and international relations: preferences, information, and empirical evidence*, P. Allan & C. Schmidt (eds). Cheltenham, England: Edward Elgar.
Nilsson, S. & D. Pitt 1994. *Protecting the atmosphere: the Climate Change Convention and its context*. London: Earthscan.

Oberthür, S. 1993. *Politik im Triebhaus: die Entstehung des internationalen Klimaschutzregimes*. Berlin: Edition Sigma.
Olson, M. & R. Zeckhauser 1966. An economic theory of alliances. *Review of Economics and Statistics* **48**, 266–79.
Ordeshook, P. C. 1986. *Game theory and political theory*. Cambridge, Massachusetts: Cambridge University Press.
Osgood, C. E. 1979. GRIT for MBFR: a proposal for unfreezing force-level postures in Europe. *Peace Research Reviews* **8**, 77–92.
Ostrom, E. 1990. *Governing the commons: the evolution of institutions for collective action*. Cambridge: Cambridge University Press.
Oye, K. 1986. *Explaining cooperation under anarchy: hypothesis and strategies*. Princeton, New Jersey: Princeton University Press.

Parson, E. A. 1996a. *A global climate-change policy exercise: results of a test run*. IIASA Working Paper WP-96-90, International Institute for Applied Systems Analysis, Laxenburg, Austria.
—— 1996b. What can you learn from a game? In *Wise choices: decisions, games, and negotiations*, R. J. Zeckhauser, R. L. Keeney, J. K. Sebenius (eds). Boston: Harvard Business School Press.
—— 1996c. *How should we study global environmental problems? A plea for unconventional methods of assessment and analysis*. IIASA Working Paper WP-96-157, International Institute for Applied Systems Analysis, Laxenburg, Austria.
Paterson, M. & M. Grubb 1992. The politics of climate. *International Affairs* **68**, 293–310.
Paxson, E. W. 1963. *War gaming report*. RM-3489, Santa Monica, California: Rand Corporation.
Peterson, J. & H. Ward 1995. Coalitional stability and the new multidimensional politics of security: a rational choice argument for US–EU cooperation. *European Journal of International Relations* **1**(2), 131–56.
Powell, R. 1991. Absolute and relative gains in international relations theory. *American Political Science Review* **85**, 1303–320.
Putnam, R. D. 1988. Diplomacy and domestic politics: the logic of two-level games. *International Organization* **42**(3), 427–60.

Raiffa, H. 1982. *The art and science of negotiation*. Cambridge, Massachusetts: Harvard University Press.

Schelling, T. C. 1960. *The strategy of conflict*. Cambridge, Massachusetts: Harvard University Press.
—— 1964. An uninhibited pitch for crisis games. Internal Research Memorandum, RAND Corporation, circulated Summer 1964 and published in *Crisis games 27 years later: plus c'est déjà vu*, P-7719, Rand Corporation, Santa Monica, California.
—— 1987. The role of war games and exercises. In *Managing nuclear operations*, A. B. Carter,

J. D. Steinbruner, C. A. Zraket (eds). Washington DC: Brookings Institution.
Sebenius, J. K. 1991. Designing negotiations towards a new regime: the case of global warming. *International Security* **15**, 110–48.
——1992. Challenging conventional explanations of international cooperation: negotiation analysis and the case of epistemic communities. *International Organization* **46**, 323–65.
——1994. Towards a winning climate coalition. In *Negotiating climate change: the inside story of the Rio Convention*, I. Mintzer & J. A. Leonard (eds). Cambridge: Cambridge University Press.
Sen, A. 1969. A games-theoretic analysis of collectivism. In *Growth and choice*, T. Majumdar (ed.). Oxford: Oxford University Press.
Shepsle, K. A. & B. R. Weingast 1975. Structure-induced equilibrium and legislative choice. *Public Choice* **23**, 503–519.
Simon, H. A. 1982. *Models of bounded rationality*. Cambridge, Massachusetts: MIT Press.
Sjostedt, G. 1993. Special and typical attributes of international environmental negotiations. In *International environmental negotiations: process, issues and contexts*, G. Sjostedt et al. (eds). Stockholm: Swedish Council for Planning and Coordination of Research and the Stockholm Institute of International Affairs.
Smoker, P. 1972. International relations simulations. In *Simulation in the social and administrative sciences*, H. Guetzkow, P. Kotler, R. L. Schultz (eds). Englewood Cliffs, New Jersey: Prentice-Hall.
——1981. The international process simulation. See Guetzkow & Valadez (1981).
Snidal, D. 1985. The limits of hegemonic stability theory. *International Organization* **39**(4), 579–614.
——1986. *The game theory of politics*. See Oye (1986).
——1991. Relative gains and the pattern of international cooperation. *American Political Science Review* **85**, 701–727.
Sonntag, N. C. 1986. Commentary. In *Sustainable development of the biosphere*, W. C. Clark & R. Munn (eds). Cambridge: Cambridge University Press.
Soroos, M. S. 1994. Global change, environmental security, and the Prisoner's Dilemma. *Journal of Peace Research* **31**, 317–32.
Stein, A. 1982. Coordination and collaboration: regimes in an anarchic world. *International Organization* **36**, 299–324.
Svedin, U. & B. Aniansson, (eds) 1987. *Surprising futures: notes from an international workshop on long-term world development*. Report 87-1, Stockholm, Sweden: Swedish Council for Planning and Coordination of Research.

Taylor, M. 1987. *The possibility of cooperation*. Cambridge: Cambridge University Press.
Taylor, M. & H. Ward 1982. Chickens, whales and lumpy public goods: alternative models of public-goods provision. *Political Studies* **30**, 350–70.
Toth, F. L. 1988a. Policy exercises: objectives and design elements. *Simulation and Games* **19**(3), 235–55.
——1988b. Policy exercises: procedures and implementation. *Simulation and Games* **19**(3), 256–76.
——1994. Practice and progress in integrated assessments of climate change: a review. In *Integrative assessment of mitigation, impacts, and adaptation to climate change*, N. Nakicenovic, W. D. Nordhaus, R. Richels, F. L. Toth (eds). Proceedings of a workshop held on 13–15 October 1993. Laxenburg, Austria: IIASA.

Valadez, J. 1981. Generational developments in modeling. See Guetzkow & Valadez (1981).
Vries, H. J. M. de 1995. SusClime: a simulation game on population and development in a resource- and climate-constrained two-country world. GLOBO Report Series 11, National Institute of Public Health and Environmental Protection [RIVM], Bilthoven, the Netherlands.
Vries, H. J. M. de, T. Fiddaman, R. Janssen. 1993. Outline for a Global Environmental Strategic

Planning Exercise (GESPE Project). GLOBO Report Series 2, National Institute of Public Health and Environmental Protection [RIVM], Bilthoven, the Netherlands.

Waltz, K. N. 1979. *Theory of international politics*. London: Addison–Wesley.
Ward, H. 1989. Testing the waters: taking risks to gain reassurance in public goods games. *Journal of Conflict Resolution* **33**, 274–308.
——1993. Game theory and the politics of the global commons. *Journal of Conflict Resolution* **37**, 203–235.
Weale, A. 1992. *The new politics of pollution*. Manchester: Manchester University Press.
WMO [World Meteorological Organization] 1986. *Report of the international conference on the assessment of the role of carbon dioxide and of other greenhouse gases in climate variations and associated impacts*. Conference Report WMO 661, World Meteorological Organization, Villach, Austria.

Young, O. R. 1989. *International cooperation: building regimes for natural resources and the environment*. Ithaca, New York: Cornell University Press.
——1991. Political leadership and regime formation: on the development of institutions in international society. *International Organization* **45**, 281–308.