



Achieving Cooperation under Anarchy: Strategies and Institutions

Robert Axelrod; Robert O. Keohane

World Politics, Vol. 38, No. 1. (Oct., 1985), pp. 226-254.

Stable URL:

<http://links.jstor.org/sici?sici=0043-8871%28198510%2938%3A1%3C226%3AACUASA%3E2.0.CO%3B2-A>

World Politics is currently published by The Johns Hopkins University Press.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/jhup.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

The JSTOR Archive is a trusted digital repository providing for long-term preservation and access to leading academic journals and scholarly literature from around the world. The Archive is supported by libraries, scholarly societies, publishers, and foundations. It is an initiative of JSTOR, a not-for-profit organization with a mission to help the scholarly community take advantage of advances in technology. For more information regarding JSTOR, please contact support@jstor.org.

ACHIEVING COOPERATION UNDER ANARCHY: Strategies and Institutions

By ROBERT AXELROD and ROBERT O. KEOHANE*

ACHIEVING cooperation is difficult in world politics. There is no common government to enforce rules, and by the standards of domestic society, international institutions are weak. Cheating and deception are endemic. Yet, as the articles in this symposium have shown, cooperation is sometimes attained. World politics is not a homogeneous state of war: cooperation varies among issues and over time.

Before trying to draw conclusions about the factors that promote cooperation under anarchy, let us recall the definitions of these key terms. Cooperation is not equivalent to harmony. Harmony requires complete identity of interests, but cooperation can only take place in situations that contain a mixture of conflicting and complementary interests. In such situations, cooperation occurs when actors adjust their behavior to the actual or anticipated preferences of others. Cooperation, thus defined, is not necessarily good from a moral point of view.

Anarchy also needs to be defined clearly. As used here, the term refers to a lack of common government in world politics, not to a denial that an international society—albeit a fragmented one—exists. Clearly, many international relationships continue over time, and engender stable expectations about behavior. To say that world politics is anarchic does not imply that it entirely lacks organization. Relationships among actors may be carefully structured in some issue-areas, even though they remain loose in others. Likewise, some issues may be closely linked through the operation of institutions while the boundaries of other issues, as well as the norms and principles to be followed, are subject to dispute. Anarchy, defined as lack of common government, remains a constant; but the degree to which interactions are structured, and the means by which they are structured, vary.

It has often been noted that military-security issues display more of the characteristics associated with anarchy than do political-economic

* We would like to thank the other authors in this project for their helpful suggestions. Robert Axelrod gratefully acknowledges the financial support of the National Science Foundation and the Harry Frank Guggenheim Foundation.

ones. Charles Lipson, for instance, has recently observed that political-economic relationships are typically more institutionalized than military-security ones.¹ This does not mean, however, that analysis of these two sets of issues requires two separate analytical frameworks. Indeed, one of the major purposes of the present collection is to show that a single framework can throw light on both.

The case studies in this symposium have shown that the three dimensions discussed in the introduction—mutuality of interest, the shadow of the future, and the number of players—help us to understand the success and failure of attempts at cooperation in both military-security and political-economic relations. Section I of this essay synthesizes some of the findings of these case studies, and thereby helps to specify some of the most important ways in which these three factors affect world politics. It deals with issues in isolation from one another, as separate games or as a series of games, in order to clarify some basic analytic points. In this section, we follow the lead of game theorists, who have tried to avoid complicating their models with extraneous material in order to reach interesting conclusions. If the problem is a small event, such as a duel between two airplanes, our analysis of it may not depend on knowledge of the context (e.g., the purpose of the war). If the issue is of very high salience to participants, such as the 1914 crisis or the Cuban missile crisis, the extraneous issues (such as tariffs, or pollution of the Caribbean) may be so insignificant that they can be ignored. Either way, the strategy of focusing only on the central interaction is clearly justified.

Yet if the issue is neither isolated nor all-consuming, the context within which it takes place may have a decisive impact on its politics and its outcomes. As the case studies illustrate, world politics includes a rich variety of contexts. Issues arise against distinctive backgrounds of past experience; they are linked to other issues being dealt with simultaneously by the same actors; and they are viewed by participants through the prisms of their expectations about the future. To ignore the effects of context would be to overlook many of the most interesting questions raised by a game-theoretic perspective on the problem of cooperation.

In Section II, we therefore consider the context of issues; in so doing, we move outward from the three dimensions on which this collection focuses toward broader considerations, including linkages among issues, multilevel games, complications encountered by strategies of reciprocity

¹ Lipson, "International Cooperation in Economic and Security Affairs," *World Politics* 37 (October 1984), 1-23.

in complex situations, and the role of international institutions. Analysis of the context of games leads us to regard context as malleable: not only can actors in world politics pursue different strategies within an established context of interaction, they may also seek to alter that context through building institutions embodying particular principles, norms, rules, or procedures for the conduct of international relations. In the conclusion, we will argue that a contextual approach to strategy—by leading us to see the importance of international institutions—helps us to forge necessary links between game-theoretic arguments and theories about international regimes.

I. THE EFFECTS OF STRUCTURE ON COOPERATION

Three situational dimensions affect the propensity of actors to cooperate: mutuality of interest, the shadow of the future, and the number of actors.

A. PAYOFF STRUCTURE: MUTUAL AND CONFLICTING PREFERENCES

It is well established that the payoff structure for a game affects the level of cooperation. For comparisons within a given type of game, this idea was first formalized by Axelrod, who established a measure of conflict of interest for specific games, including Prisoners' Dilemma.² Experimental evidence demonstrated that the greater the conflict of interest between the players, the greater the likelihood that the players would in fact choose to defect. Jervis has elaborated on these theories and shown that different types of games, such as Stag Hunt and Chicken, have different potentials for cooperation.³ He has also applied his strategic analysis to historical and contemporary problems related to the security dilemma. His work clearly indicates that international cooperation is much easier to achieve in some game settings than in others.

Payoff structures often depend on events that take place outside of the control of the actors. The economic depressions of 1873-1896 and of the early 1930s stimulated demands for protection by firms and individuals in distress, and therefore reduced the incentives of governments to cooperate with one another. The weakness and vacillation of the British and French governments before 1939 reduced the potential value

² Robert Axelrod, "Conflict of Interest: An Axiomatic Approach," *Journal of Conflict Resolution* 11 (March 1967), 87-99; and *Conflict of Interest: A Theory of Divergent Goals with Applications to Politics* (Chicago: Markham, 1970).

³ Robert Jervis, "Cooperation under the Security Dilemma," *World Politics* 30 (January 1978), 167-214.

of anti-German alliances with those countries for the Soviet Union, making a Nazi-Soviet pact seem relatively more attractive.

This is obvious enough. Slightly less obvious is another point about mutuality of interests: the payoff structure that determines mutuality of interests is not based simply upon objective factors, but is grounded upon the actors' perceptions of their own interests. Perceptions define interests. Therefore, to understand the degree of mutuality of interests (or to enhance this mutuality) we must understand the process by which interests are perceived and preferences determined.

One way to understand this process is to see it as involving a change in payoffs, so that a game such as Prisoners' Dilemma becomes either more or less conflictual. To start with, Prisoners' Dilemma is a game in which both players have an incentive to defect no matter whether the other player cooperates or defects. If the other player *cooperates*, the first player prefers to defect: $DC > CC$. On the other hand, if the other player *defects*, the first player still prefers to defect: $DD > CD$. The dilemma is that, if both defect, both do worse than if both had cooperated: $CC > DD$. Thus, Prisoners' Dilemma has a preference ordering for both players of $DC > CC > DD > CD$.⁴

Now consider a shift in the preferences of both players, so that mutual cooperation is preferred to unilateral defection. This makes the preference ordering $CC > DC > DD > CD$, which is a less conflictual game called Stag Hunt.

Jervis's study of the shift from balance-of-power systems to concerts suggests that after world wars, the payoff matrix for the victors may temporarily be one of Stag Hunt: fighting together results in a short-lived preference for staying together. After a war against a hegemonic power, the other great powers often perceive a mutual interest in continuing to work together in order to ensure that the defeated would-be hegemon does not rise again. They may even feel empathy for one another, and take an interest in each other's welfare. These perceptions seem to have substantial momentum, both among the mass public and in the bureaucracy. Yet, the cooperation that ensues is subject to fairly easy disruption. As recovery from the war proceeds, one or both parties may come to value cooperation less and relative gains more. And if one side believes that its counterpart prefers to defect, its own preference will shift to defection in order to avoid the worst payoff, CD.

Actors can also move from Prisoners' Dilemma to more conflictual

⁴The definition of Prisoners' Dilemma also includes one additional restriction: $CC > (DC + CD)/2$. This is to ensure that it is better to have mutual cooperation than to have an even chance of being the exploiter or the exploited.

games. If both players come to believe that mutual cooperation is worse than mutual defection, the game becomes Deadlock, with both sides having preferences of $DC > DD > CC > CD$. Since the dominant strategy of each player is to defect regardless of what the other does, the likely outcome is DD. Players in Deadlock, unlike those in Prisoners' Dilemma, will not benefit from repeated plays since mutual cooperation is not preferred to mutual defection.

Kenneth Oye provides a fine example of the movement from Prisoners' Dilemma to Deadlock in his essay on monetary diplomacy in the 1930s, in this collection. Shifts in beliefs, not only about international regimes, but particularly about desirable economic policy, led leaders such as Franklin D. Roosevelt to prefer unilateral, uncoordinated action to international cooperation on the terms that appeared feasible. Oye argues that the early 1930s do not mark a failure of coordination where common interests existed (as in Prisoners' Dilemma); rather, they indicate the decay of these common interests, as perceived by participants. In their essay in the present collection, Downs, Rocke, and Siverson argue that arms races are often games of Deadlock rather than Prisoners' Dilemma, making them much more difficult to resolve.

Beliefs are as important in the military area as in economics. Consider, for example, Van Evera's study of the beliefs leading to World War I. By 1914, what Van Evera labels "the cult of the offensive" was universally accepted in the major European countries. It was a congenial doctrine for military elites everywhere, since it magnified the role of the military and reduced that of the diplomats. It also happened to be disastrously wrong, since its adherents failed to appreciate the overwhelming advantage that recent technological change had given to the defensive (in what was soon to become trench warfare), and overlooked the experiences of the American Civil War and the Russo-Japanese War.

Gripped by this cult of the offensive, European leaders sought to gain safer borders by expanding national territories, and took more seriously the possibility of successful aggressive war; hence Germany and (to a lesser extent) other European powers adopted expansionist policies that brought them into collision with one another. European leaders also felt greater compulsion to mobilize and strike first in a crisis, since the penalty of moving late would be greater in an offense-dominant world; this compulsion then fueled the spiral of mobilization and counter-mobilization that drove the July 1914 crisis out of control. Had Europeans recognized the actual power of the defense, expansionism would have lost much of its appeal, and the compulsion to mobilize and counter-mobilize would have diminished. Put differently, the European payoff

structure actually would have rewarded cooperation; but Europeans perceived a payoff structure that rewarded *non*cooperation, and responded accordingly. Beliefs, not realities, governed conduct.

The case of 1914 also illustrates a point made above: subjective interpretations by one side become objective reality for the other side. When a European state adopted expansionist policies, those nearby found themselves with an expansionist neighbor, and had to adjust accordingly. For instance, Germany's expansionism, though largely based on illusions, led to a genuine change in Russia's environment. Russia adopted its inflexible war plan (which required mobilization against Germany as well as against Austria) partly because the Russians feared that Germany would strike into Russia's northern territories once the Russian armies were embroiled with Austria. Thus the Russian calculus was importantly affected by Russia's image of German intent, and Russia was driven to bellicose measures by fear of German bellicosity. German expansionism was premised largely on illusions, but for Russia this expansionism was a real danger that required a response.

This discussion of payoff structures should make it clear that the contributors to this volume do not assume that Prisoners' Dilemmas are typical of world politics. More powerful actors often face less powerful ones, yielding asymmetric payoff matrices. Furthermore, even symmetrical games can take a variety of forms, as illustrated by Stag Hunt, Chicken, and Deadlock. What is important for our purposes is not to focus exclusively on Prisoners' Dilemma *per se*, but to emphasize the fundamental problem that it (along with Stag Hunt and Chicken) illustrates. In these games, myopic pursuit of self-interest can be disastrous. Yet both sides can potentially benefit from cooperation—if they can only achieve it. Thus, choices of strategies and variations in institutions are particularly important, and the scope for the exercise of intelligence is considerable.

Our review of payoff structures also illustrates one of the major themes of this collection of essays: that political-economic and military-security issues can be analyzed with the same analytical framework. Admittedly, economic issues usually seem to exhibit less conflictual payoff structures than do those of military security. Coordination among bankers, as described by Lipson, has been more extensive and successful than most arms control negotiations, as analyzed by Downs and his colleagues; and the patterns of trade conflict and cooperation described by Conybear are hardly as conflictual as Van Evera's story of World War I. On the other hand, the great power concerts discussed by Jervis, as well as several of the arms control negotiations, were more cooperative than the

trade and monetary measures of 1930-1933 delineated in Oye's essay. And postwar economic relations between the United States and Japan have been more conflictual than military-security relations. As an empirical matter, military issues may more often have payoff structures involving a great deal of conflict of interest; but there is no theoretical reason to believe that this must always be the case.⁵

B. THE SHADOW OF THE FUTURE

In Prisoners' Dilemma, concern about the future helps to promote cooperation. The more future payoffs are valued relative to current payoffs, the less the incentive to defect today—since the other side is likely to retaliate tomorrow.⁶ The cases discussed in the present essays support this argument, and identify specific factors that help to make the shadow of the future an effective promoter of cooperation. These factors include:

1. long time horizons;
2. regularity of stakes;
3. reliability of information about the others' actions;
4. quick feedback about changes in the others' actions.

The dimension of the shadow of the future seems to differentiate military from economic issues more sharply than does the dimension of payoffs. Indeed, its four components can be used to analyze some of the reasons why issues of international political economy may be settled more cooperatively than issues of international security, even when the underlying payoff matrices are similar—for example, when Prisoners' Dilemma applies. Most important is a combination of the first two factors: long time horizons and regularity of stakes. In economic relations, actors have to expect that their relationships will continue over an indefinite period of time; that is, the games they play with each other will be iterated. Typically, neither side in an economic interaction can eliminate the other, or change the nature of the game decisively in a single move. In security affairs, by contrast, the possibility of a successful preemptive war can sometimes be a tempting occasion for the rational timing of surprise.⁷ Another way to put this is that, in the international political economy, retaliation for defection will almost always be possible;

⁵ For an earlier discussion of contemporary events, using a common analytical framework to examine both economic and security relations, see Oye, "The Domain of Choice," in Kenneth A. Oye, Donald Rothchild, and Robert J. Lieber, eds., *Eagle Entangled: U.S. Foreign Policy in a Complex World* (New York: Longman, 1979), 3-33.

⁶ Robert Axelrod, *The Evolution of Cooperation* (New York: Basic Books, 1984).

⁷ Robert Axelrod, "The Rational Timing of Surprise," *World Politics* 31 (January 1979), 228-46.

therefore a rational player, considering defection, has to consider its probability and its potential consequences. In security affairs, it may be possible to limit or destroy the opponent's capacity for effective retaliation.

To illustrate this point, let us compare the case of 1914 with contemporary international debt negotiations. In 1914, some Germans, imbued with the cult of the offensive, thought that a continental war would permanently solve Germany's security problems by restructuring power and territorial relations in Europe. For these German leaders, the temptation to defect was huge, largely because the shadow of the future seemed so small. Indeed, it seemed that future retaliation could be prevented, or rendered ineffective, by decisive German action. Moreover, in the opening move of a war the stakes would be far greater than usual because of the value of preempting before the other side was fully mobilized. This perceived irregularity in the stakes further undercut the potential for sustained cooperation based upon reciprocity.

By contrast, contemporary negotiations among banks, and between banks and debtor countries, are heavily affected by the shadow of the future. That is not to say that the stakes of each game are the same; indeed, there are great discontinuities since deadlines for rescheduling take on importance for regulators, banks, and the reputations of borrowers. But the banks know that they will be dealing both with the debtor countries and with one another again and again. Continuing interbank relationships imply, as Lipson points out, that small banks will think twice before doublecrossing large banks by refusing to participate in rescheduling. This is particularly true if the small banks are closely tied, in a variety of ways, to the large banks. Continuing relations between banks and debtor countries give the banks incentives to cooperate with the debtor countries, not merely in order to facilitate debt servicing on loans already made, but to stay in their good graces—looking toward a more prosperous future. The fact that Argentina, Brazil, and Mexico are so large, and are perceived to be potentially wealthy, is a significant bargaining asset for them now, since it increases the banks' expected profits from future lending, and therefore enlarges the shadow of the future. Indeed, if these governments could credibly promise to favor, in the future, banks that help them now, and to punish or ignore those that defect in these critical times, they could further improve their bargaining positions; but, as sovereign governments whose leaders will be different in the future, they cannot effectively do so.

Reliability of information about the others' actions and promptness of feedback are also important in affecting the shadow of the future,

although they do not seem to differentiate military-security from political-economic issues so clearly. Because of the absence of military secrecy, actors may sometimes have more reliable information on political-economic than on military-security issues. Banks thrive on differential access to information, and therefore hold it closely. Furthermore, since the systemic effects of political-economic actions are often difficult to judge, and "cheating at the margin" is frequently easy, feedback between policy and results may be slow. For instance, the distribution of benefits from the Tokyo Round of trade negotiations is still a matter of conjecture and political contention rather than economic knowledge. By contrast, the superpowers publish lists of the precise number of missiles in each other's inventories, and we can assume that information about the effect of a military action by either side—short of a devastating surprise attack that would destroy command and control facilities—would be communicated almost immediately to the leaders of both states.

The length of the shadow of the future, like the character of payoff structures, is not necessarily dictated by the objective attributes of a situation. On the contrary, as we have just seen, expectations are important. International institutions may therefore be significant, since institutions embody, and affect, actors' expectations.⁸ Thus institutions can alter the extent to which governments expect their present actions to affect the behavior of others on future issues. The principles and rules of international regimes make governments concerned about precedents, increasing the likelihood that they will attempt to punish defectors. In this way, international regimes help to link the future with the present. That is as true of arms control agreements, in which willingness to make future agreements depends on others' compliance with previous arrangements, as it is in the General Agreement on Tariffs and Trade, which embodies norms and rules against which the behavior of members can be judged. By sanctioning retaliation for those who violate rules, regimes create expectations that a given violation will be treated not as an isolated case but as one in a series of interrelated actions.

C. NUMBER OF ACTORS: SANCTIONING PROBLEMS

The ability of governments to cooperate in a mixed-motive game is affected not only by the payoff structure and the shadow of the future, but also by the number of players in the game and by how their relationships are structured. Axelrod has shown that reciprocity can be an

⁸ Stephen D. Krasner, ed., *International Regimes* (Ithaca, NY: Cornell University Press, 1983); Robert O. Keohane, *After Hegemony: Cooperation and Discord in the World Political Economy* (Princeton: Princeton University Press, 1984).

effective strategy to induce cooperation among self-interested players in the iterated, bilateral Prisoners' Dilemma, where the values of each actor's options are clearly specified.⁹ However, effective reciprocity depends on three conditions: (1) players can identify defectors; (2) they are able to focus retaliation on defectors; and (3) they have sufficient long-run incentives to punish defectors. When there are many actors, these conditions are often more difficult to satisfy. In such situations, it may be impossible to identify, much less to punish, defection; even if it is possible, none of the cooperators may have an incentive to play the role of policeman. Each cooperator may seek to be a free-rider on the willingness of others to enforce the rules.

We may call the difficulty of preventing defection through decentralized retaliation the "sanctioning problem." Its first form, the inability to identify defectors, is illustrated by the terrorist bombings against American installations in Lebanon in 1983. The United States did not know, at the time the bombings took place, who was responsible. The only state that could plausibly have been held responsible was Syria; but since the Syrians denied responsibility, retaliation against Damascus could have spread and deepened the conflict without punishing the terrorist groups themselves. The issue of identifying defectors is one aspect of a fundamental problem besetting efforts to cooperate in world politics: acquiring, in a timely fashion, adequate amounts of high-quality information. In order to maintain cooperation in games that reward unreciprocated defection, such as Prisoners' Dilemma, governments must have confidence in their ability to monitor their counterparts' actions sufficiently well to enable them to respond effectively to betrayal. As Lipson has pointed out, the greater perils of betrayal (to the side that is betrayed) in military-security than in political-economic relations put more severe demands on gathering information in the former than in the latter area.¹⁰

The second form of the sanctioning problem occurs when players are unable to focus retaliation on defectors. This difficulty is illustrated by Conybeare's analysis of the Anglo-Hanse trade wars. The Hanseatic League was unable to punish English privateers for their depredations, and instead retaliated against English merchants in Hanseatic towns. This produced escalation rather than cooperation.

The third form of the sanctioning problem arises when some members of a group lack incentives to punish defectors. This obstacle to cooperation often arises where there are many actors, some of which fail

⁹ Axelrod (fn. 6).

¹⁰ Lipson (fn. 1).

to cooperate in the common effort to achieve some collective good. Oye observes that although British devaluation in 1931 hurt other countries, no single government had the incentive to devote its own resources to bring about a revision of British policy. This form of the sanctioning problem—lack of incentives to punish defectors—also arose in the debt negotiations of the 1980s. To prevent default, it was necessary to arrange rescheduling agreements involving additional bank lending. Smaller banks were tempted to refuse to provide new funds. Only the fact that the large banks had strong incentives to put pressure on smaller ones to ante up prevented rescheduling agreements from unravelling “like a cheap sweater.”

When sanctioning problems are severe, cooperation is in danger of collapsing. One way to bolster it is to restructure the situation so that sanctioning becomes more feasible. Sometimes this is done unilaterally. Oye points out that external benefits or costs may be “privatizable”; that is, changes can be made in the situation so that the benefits and costs of one’s actions are directed specifically at those with whom one has negotiated. He argues that in the early 1930s Britain eventually succeeded in privatizing its international currency relationships by adopting exchange controls and attaching conditions, negotiated bilaterally, to new loans. This transformation of the game permitted a modest revival of international lending, based not on open access to British capital markets but on bilateral reciprocity.

As our examples indicate, sanctioning problems can occur both in the international political economy and on military-security issues. They tend to be more severe on military-security than on political-economy issues, due to the high costs of punishing defections, the difficulties of monitoring behavior, and the stringent demands for information that are imposed when successful defection can dramatically shorten the shadow of the future. But since sanctioning problems occur on both types of issues, issue-area alone cannot account for their incidence or severity. To explain the incidence and severity of sanctioning problems, we need to focus on the conditions that determine whether defection can be prevented through decentralized retaliation: the ease of identifying sources of action, the ability of governments to focus retaliation or reward on particular targets, and the incentives that exist for members of a group to punish defectors.

While the likelihood that these problems will arise may be enhanced by an increase in the number of actors involved, difficulties may also appear on issues that seem at first glance to be strictly bilateral. Consider, for instance, the example of 1914. In the Balkan crisis, Austria sought

to impose sanctions against Serbia for its support of revolutionaries who tried to destroy the ethnically heterogeneous Austro-Hungarian empire. But sanctions against Serbia implied punishment for Russia, Serbia's ally, since Russian leaders were averse to accepting another Balkan setback. Russian mobilization, however, could not be directed solely against Austria, since Russia only had plans for general mobilization.¹¹ Thus, neither Austria nor Russia was able to focus retaliation on the defector; the actions of both helped to spread rather than to contain the crisis. With more clever and moderate leadership, Austria might have found a way to punish Serbia without threatening Russia. And a detailed plan for mobilization only against Austria could have provided Russia with a more precisely directed measure to retaliate against Austria's ultimatum to Serbia.

Privatization is not the only way to maintain cooperation. Moreover, as some of our examples indicate, it can be difficult to achieve. Another way to resolve sanctioning problems is to construct international regimes to provide standards against which actions can be measured, and to assign responsibility for applying sanctions. Regimes provide information about actors' compliance; they facilitate the development and maintenance of reputations; they can be incorporated into actors' rules of thumb for responding to others' actions; and they may even apportion responsibility for decentralized enforcement of rules.¹²

Charles Lipson's discussion of the international lending regime that has been constructed by bankers reveals how regimes can promote cooperation even when there are many actors, no dominant power, and no world central bank. Creditor committees were established under the leadership of large money-center banks. Each money-center bank then took responsibility for a number of relatively large regional banks, which in turn were assigned similar responsibilities for smaller banks.¹³ As a result, a hierarchy of banks was created, isolating smaller banks from one another and establishing responsibility for enforcing sanctions. Small banks displaying tendencies toward defection were threatened with being outside the flow of information in the future and, implicitly, with not being offered participation in lucrative future loans. This informal hierarchy, of course, was reinforced by the presence of the U.S. Federal Reserve System looming in the background: stories, whether apocryphal or not, of small bankers being told to "cough up" by high officials of

¹¹ Robert E. Osgood and Robert W. Tucker, *Force, Order and Justice* (Baltimore: Johns Hopkins University Press, 1967), esp. chap. 2, "The Expansion of Force."

¹² Keohane (fn. 8), 49-132.

¹³ Lipson, "Bankers' Dilemmas," in this collection, 200-225.

the Fed circulated in banking circles during the early 1980s. It would have taken a bold president of a small bank to ignore both the banking hierarchy and the danger of arousing the Fed's wrath by not participating in a rescheduling.

This reference to the role of institutions in transforming N-person games into collections of two-person games suggests once again the importance of the context within which games are played. In isolation, the basic concepts discussed in the introduction—payoff structures, iteration, and the number of players—provide only a framework for analysis. They take on greater significance, as well as complexity, when they are viewed within the broader context of other issues, other games, and the institutions that affect the course of world politics. We now turn to the question of how the context of interaction affects political behavior and outcomes.

II. THE CONTEXT OF INTERACTION

Whether cooperation can take place without central guidance depends not merely on the three game-theoretic dimensions we have emphasized so far, but also on the context within which interaction takes place. Context may, of course, mean many different things. Any interaction takes place within the context of norms that are shared, often implicitly, by the participants. John Ruggie has written of the "deep structure" of sovereignty in world politics,¹⁴ and also of the way in which shifting values and norms of state intervention in society—the emergence and legitimation of the welfare state—affected the world political economy between 1914 and 1945. International political-economic bargaining was fundamentally changed by the shift, during this period, from laissez-faire liberalism as a norm to what Ruggie calls "embedded liberalism."¹⁵

Interactions also take place within the context of institutions. Robert Keohane has argued elsewhere that even if one adopts the assumption that states are rational and self-interested actors, institutions can be shown to be important in world politics.¹⁶ Institutions alter the payoff structures facing actors, they may lengthen the shadow of the future,

¹⁴ John G. Ruggie, "Continuity and Transformation in the World Polity: Toward a Neorealist Synthesis," *World Politics* 35 (January 1983), 261-85.

¹⁵ John G. Ruggie, "International Regimes, Transactions and Change: Embedded Liberalism in the Postwar Economic Order," *International Organization* 36 (Spring 1982), 379-416, reprinted in Krasner (fn. 8), 195-231; Fred Hirsch, "The Ideological Underlay of Inflation," in John Goldthorpe and Fred Hirsch, eds., *The Political Economy of Inflation* (London: Martin Robertson, 1978), 263-84.

¹⁶ Keohane (fn. 8).

and they may enable N-person games to be broken down into games with smaller numbers of actors.

Using the game-theoretic perspective of this symposium, another way of looking at context may be especially revealing. This aspect has to do with what we call multilevel games. In such situations, different games affect one another, so that their outcomes become mutually contingent. Three such situations are particularly important for world politics: issue-linkage, domestic-international connections, and incompatibilities between games among different sets of actors. After considering these situations, we will turn to the implications of these multilevel games for the efficacy of a strategy of reciprocity in fostering cooperation.

A. MULTILEVEL GAMES

Issue-linkage. Most issues are linked to other issues. This means that games being played on different issues—different “chessboards,” in Stanley Hoffmann’s phrase¹⁷—affect one another. Connections between games become important when issues are linked.

Issue-linkage in this sense involves attempts to gain additional bargaining leverage by making one’s own behavior on a given issue contingent on others’ actions toward other issues.¹⁸ Issue-linkage may be employed by powerful states seeking to use resources from one issue-area to affect the behavior of others elsewhere; or it may be employed by outsiders, attempting to break into what could otherwise be a closed game. Linkage can be beneficial to both sides in a negotiation, and can facilitate agreements that might not otherwise be possible.¹⁹ Actors’ resources may differ, so that it makes sense to trade one for the other: the United States, for instance, may provide economic aid to Egypt in exchange for Egyptian support for American policy in the Middle East. Furthermore, different players may have preferences of different intensities: thus, in a log-rolling game, each party trades its “vote,” or policy position, on an issue it values less highly for the other’s vote on one it values more highly.

The outstanding example of a successful bargaining linkage in our

¹⁷ Stanley Hoffmann, “International Organization and the International System,” *International Organization* 24 (Summer 1970), 389-413.

¹⁸ Ernst B. Haas refers to this as “tactical” issue-linkage, contrasting it with “substantive” issue-linkage resulting from causal knowledge. See Haas, “Why Collaborate? Issue-linkage and International Regimes,” *World Politics* 32 (April 1980), 357-405, at 372. For a sophisticated analysis of tactical issue-linkage, see Michael McGinnis, “Issue Linkage and the Evolution of International Cooperation,” *Journal of Conflict Resolution*, forthcoming.

¹⁹ Robert E. Tollison and Thomas D. Willett, “An Economic Theory of Mutually Advantageous Issue Linkage in International Negotiations,” *International Organization* 33 (Fall 1979), 425-49.

case studies is that of the Washington Naval Treaty of 1922. As Downs, Rocke, and Siverson show, these arms control negotiations were successful in part because they linked bargaining over arms with bargaining over other issues. As part of an agreement to limit battleship construction, Japan gave Britain and the United States guarantees regarding trade in China and limitations of fortification on certain Pacific islands; Japan received legal recognition of its right to certain territory taken from Germany after World War I. Bringing these issues into the negotiations to limit the building of battleships helped to make cooperation possible, not only on these specific issues but on the whole package.

Of course, not all issue-linkages promote agreement, any more than each exercise of power can be expected to lead to cooperation. Oye has distinguished between “backscratching,” which he regards as welfare-enhancing, and “blackmailing,” which may reduce welfare levels.²⁰ The “backscratcher” merely offers, in return for compensation, to refrain from acting in what would otherwise be its own best interest. For instance, a debtor country, unable to make its payments on time without facing severe hardship or political revolution, may offer to continue servicing its debts only if compensated with new loans and an easier payment schedule. If this offer is rejected, the debtor does what it would have done without the offer: it defaults.

Backscratching entails a promise. Blackmailing, by contrast, implies a threat. As Schelling has pointed out, “the difference is that a promise is costly when it succeeds, and a threat is costly when it fails.”²¹ Blackmailers threaten to act against their own interests unless compensated. Thus, a debtor country that would be hurt by defaulting may nevertheless threaten to do so unless compensation is offered. This threat, if carried out, would leave both the debtor (the blackmailer, in this case) and its creditors worse off than if it had merely acted in its own interest without bargaining at all. If the blackmailing strategy works, on the other hand, the effect will be to transfer resources from the creditors to the debtor, an action that will not necessarily improve overall welfare.

Although it may be difficult to differentiate between backscratching and blackmailing in practice, the distinction helps us to recognize that issue-linkages have dangers as well as opportunities. One side may demand so much of the other in other areas that cooperation will not take place even in the area of shared interests. This accusation is fre-

²⁰ Oye (fn. 5).

²¹ Thomas C. Schelling, *The Strategy of Conflict* (New York: Oxford University Press, 1960), 177.

quently made against Henry Kissinger's version of linkage. Kissinger insisted that the Soviets exercise great restraint in the Third World in return for American cooperation on arms control.²² In Oye's terms, Kissinger was trying to "blackmail" the Soviets by threatening to act against the United States' own interests (delay arms control) unless the Soviets compensated the United States with unilateral restraint.²³

The most intriguing point about linkage that is highlighted by the case studies is the existence of what could be called "contextual" issue-linkage. In such a situation, a given bargain is placed within the context of a more important long-term relationship in such a way that the long-term relationship affects the outcome of the particular bargaining process. Two cases of contextual issue-linkage show that this form can often work to reduce conflict even without affecting the preferences of the participants on the specific issues being discussed. Oye notes that in 1936, the United States, Britain, and France were able to reach an agreement on international monetary reform because of the common security concern over a rising Nazi Germany. And as Downs and his colleagues point out, by far the most important cause of cooperation in arms races that ended peacefully has been the activity of a third power. For example, the Anglo-French naval arms race of 1852-1853 was resolved when the two states formed an alliance in order to fight the Russians in the Crimean War.

International relations and domestic politics. Similar analytic questions arise in considering connections between international relations and domestic politics. Arms control negotiations involve not merely bargaining between governments, but within societies as well; the Carter administration was able to resolve the SALT II game with the Soviet Union, but not with the U.S. Senate. Trade issues typically also involve both international and domestic games. In the Tokyo Round, the same Carter administration—with a different responsible party, Robert Strauss—was able to mesh international and domestic games, playing them simultaneously rather than sequentially (international first), as had been done on some issues in the Kennedy Round a decade earlier. The result in this case was that the Tokyo Round trade agreements with

²² George W. Breslauer, "Why Détente Failed: An Interpretation," in Alexander L. George and others, *Managing U.S.-Soviet Rivalry: Problems of Crisis Prevention* (Boulder, CO: Westview Press, 1983), 319-40; John L. Gaddis, "The Rise, Fall and Future of Détente," *Foreign Affairs* 62 (Winter 1983/84), 354-77; Stanley Hoffmann, "Détente," in Joseph S. Nye, ed., *The Making of America's Soviet Policy* (New Haven: Yale University Press for the Council on Foreign Relations, 1984), 231-64.

²³ Oye (fn. 5), 17.

other countries were all ratified overwhelmingly by Congress, in contrast to the rejection of some of the international agreements made in the Kennedy Round.²⁴

Such domestic-international connections are commonplace. Frequently, the incentives provided by domestic bargaining games inhibit effective foreign policy and may exacerbate international conflict. A well-known case is that of American decision making during the early months of the Korean War. General MacArthur was such a formidable figure in American politics that even his military superiors were reluctant to challenge his judgment in marching toward the Yalu River in the fall of 1950; yet this maneuver was so questionable that, if it had not been for the domestic political games taking place, serious reservations would have been expressed in the Pentagon and the White House.²⁵

Another type of domestic-international linkage is discussed by Coynbeare in this collection. During the 15th century, the Hanseatic League responded to naval setbacks at the hands of Britain by financing and equipping Edward IV, who, upon defeating the Lancastrians in the War of the Roses, signed a treaty that was one-sidedly favorable to the Hanse's trading interests. By intervening in British domestic politics, the Hanse was thus able to triumph despite military weakness. This technique—intervening in a domestic political game as compensation for weakness at the international level—has recently been employed in more subtle ways by small powers with strong interests in American foreign policy.²⁶

Compatibilities and incompatibilities among games. Many different games take place in world politics, involving different but overlapping sets of actors. Sometimes the existence of more than one game makes it easier to attain cooperation, but related games may also create difficulties for one another. That is, games in world politics can be compatible or incompatible with each other.

One example of a set of compatible games is provided by cooperation in international economic negotiations among the major industrialized countries. After World War II, such cooperation was facilitated by the fact that these countries were military allies. In contrast to Britain's situation in the 19th century, America's ability to persuade other major trading states to accept the rules that it preferred was greatly enhanced by the fact that in the military-political game the United States was a

²⁴ Gilbert Winham, "Robert Strauss, the MTN, and the Control of Faction," *Journal of World Trade Law* 14 (September-October 1980).

²⁵ Alexander George and Richard Smoke, *Deterrence in American Foreign Policy* (New York: Columbia University Press, 1974).

²⁶ Robert O. Keohane, "The Big Influence of Small Allies," *Foreign Policy*, No. 2 (Spring 1971), 161-82.

senior partner, rather than an adversary, of the other major actors in the world economy. To take another example: Lipson's analysis of debt negotiations suggests that the negotiating game among large banks was rendered compatible with games between large and small banks by structuring the situation so that small banks could not coordinate with each other. That is, two sets of negotiations were made compatible by precluding a third one.

The case of 1914 illustrates the problem of incompatibility among games. In non-crisis periods, loyalty within an alliance was compatible with friendly relations across alliances. But when the 1914 crisis occurred, loyalty within an alliance—such as Germany's support for Austria, Russia's for Serbia, and France's for Russia—implied defection across alliances. The increased cooperativeness of intra-alliance games destroyed broader patterns of cooperation.

In the contemporary international political economy, problems of incompatibility may also arise. For instance, negotiations on questions such as tariffs or energy policies are most likely to yield positive results for the advanced industrialized countries when only a few major players are involved in the initial negotiation. Friction with others, however, especially the less developed countries, may produce conflict on a larger scale. Or, to take a different example from the politics of international debt, close and explicit collaboration among debtor countries could, some fear, disrupt relations between debtor governments and banks in the richer countries.

The contrast between the fate of Soviet-American arms control in the 1970s and the Tokyo Round of trade negotiations illustrates the importance of multilevel games. In the face of linkages to other contentious issues, complex domestic political games, and a lack of reinforcement between political-economic and military-security games, even shared interests, a long shadow of the future, and bilateralism may be insufficient to promote cooperation. If the interaction happens to be an iterated game of Chicken, the problem is even worse because each player has a strong incentive to avoid cooperation in the short run in order to develop a reputation for firmness in the long run. Conversely, even when there are quite severe conflicts of interest, these may be overshadowed by more important mutual interests, perhaps institutionalized in organizations such as NATO. Once again, it is not sufficient to analyze a particular situation in isolation from its political context. We must also analyze the patterns of expectations, and the institutions created by human beings, within which particular negotiations are located and in the light of which they are interpreted by participants.

B. RECIPROCITY AS A STRATEGY IN MULTILEVEL GAMES

Robert Axelrod has employed computer tournaments and theoretical analysis of the iterated, two-player Prisoners' Dilemma to show that a strategy based on reciprocity—such as Tit-for-Tat—can be remarkably effective in promoting cooperation.²⁷ Even among pure egoists, cooperation can “emerge” if a small initial cluster of potential cooperators exists.

This argument suggests that governments may have incentives to practice reciprocity in a variety of situations that are characterized by mixtures of conflicting and complementary interests—that is, in certain non-zero-sum games. Evidence for this proposition is established best for the particular case of Prisoners' Dilemma. Axelrod's theory suggests that in this game a strategy based on reciprocity can yield relatively high payoffs against a variety of other strategies. Furthermore, such a strategy helps the whole community by punishing players who use uncooperative strategies. When payoff structures are those of Prisoners' Dilemma, therefore, we can expect practitioners of reciprocity to attempt to institutionalize it as a general practice, so that they will benefit from others' use of the strategy as well as their own.

As we have noted above, not every situation in which conflict or cooperation may occur can be categorized as Prisoners' Dilemma. Games such as Chicken and Stag Hunt are also significant. Evidence on these cases is not as extensive as on Prisoners' Dilemma. Yet, as Oye's introduction points out, there are good reasons to believe that reciprocity is an attractive strategy in a variety of non-zero-sum situations. The key conditions for the successful operation of reciprocity are that mutual cooperation can yield better results than mutual defection, but that temptations for defection also exist. In such situations, reciprocity may permit extensive cooperation without making cooperative participants inordinately vulnerable to exploitation by others. Furthermore, it may deter uncooperative actions.²⁸

²⁷ Axelrod (fn. 6).

²⁸ Consider the example of Stag Hunt, defined by the preference ordering of both players as $CC > DC > DD > CD$. If Player *A* is credibly committed to a strategy of reciprocity, beginning with cooperation, *B*'s incentives to cooperate are enhanced. *A*'s commitment to cooperate ensures that *B* will not be double-crossed (which would leave *B* with the worst payoff). Furthermore, *A*'s commitment to retaliate against defection ensures that any defection by *B* would lead, after the first move, not to *B*'s second-best outcome (DC), but to its third-best outcome (DD). The game of Chicken provides another appropriate case in point. In Chicken, mutual cooperation is only the second-best outcome for both players, but mutual defection is worst for both. Thus, $DC > CC > CD > DD$. A credible strategy of reciprocity by Player *A* in Chicken ensures *B* of its second-best outcome if it cooperates, and guarantees that continual defection will in the long run provide it with its worst payoff. Assuming that *B*'s shadow of the future is sufficiently long, it should respond to *A*'s strategy of reciprocity by cooperating.

It is not surprising, therefore, that reciprocity is a popular strategy for practical negotiators as well as for analysts in the laboratory. Oye's analysis of monetary politics in the 1930s reveals that Britain developed such a strategy in its relations with the Scandinavian countries. Contemporary discussions of international trade provide another case in point. U.S. officials have frequently defended reciprocity in trade relations on the grounds that pursuit of this strategy would deter discrimination against American products by other countries, and that relaxation of reciprocity would invite retaliation by others. Even observers skeptical about reciprocity often agree. In a policy-oriented article critical of current proposals that the United States should practice "aggressive reciprocity" in trade negotiations, William Cline argues that such action is rendered less effective by a high probability of foreign counter-retaliation.²⁹ In Axelrod's terms, Tit-for-Tat (which begins by cooperating and then retaliates once for each defection by the other player) discourages exploitative strategies—"aggressive reciprocity."

Thus, the applicability of Tit-for-Tat does not seem to be limited to Prisoners' Dilemma. Yet it is not a perfect strategy. In the first place, it can perpetuate conflict through an "echo effect": "if the other player defects once, Tit-for-Tat will respond with a defection, and then if the other player does the same in response, the result would be an unending echo of alternating defections."³⁰ In real-world politics as well as in the laboratory, reciprocity can lead to feuds as well as to cooperation, particularly when players have different perceptions of past outcomes.³¹ Soviet-American détente collapsed partly because each side concluded that the other was not practicing reciprocity, but was, on the contrary, taking unilateral advantage of its own restraint.³² Second, even when many shared interests exist and judgments of equivalence are not distorted, reciprocity may lead to deadlock. John W. Evans has pointed out that in tariff negotiations conducted according to the principle of reciprocity, potential concessions may become "bargaining chips" to be hoarded: "Tariffs that have no intrinsic economic value for a country that maintains them have acquired value because of the insistence of other countries on reciprocity in the bargaining process." As a result, "tariff levels may be maintained in spite of the fact that a lower level

²⁹ Cline, " 'Reciprocity': A New Approach to World Trade Policy?" Institute for International Economics, Policy Analyses in International Economics 2 (Washington: September 1982), 25.

³⁰ Axelrod (fn. 6), 176.

³¹ For an analysis of the spiral mode of conflict, see Robert Jervis, *Perception and Misperception in International Politics* (Princeton, NJ: Princeton University Press, 1976), esp. 58-113.

³² See references cited in fn. 22.

would raise the country's real income."³³ Third, when several actors negotiate separately and sequentially over issues that are substantively interdependent, subsequent bargains may call previous agreements into question by altering the value of concessions that have been made. This "issue interdependence problem" bedeviled trade negotiations under the conditional most-favored-nation clause prior to the institution of multilateral trade negotiations after World War II. Conditional most-favored-nation treatment permitted discrimination among suppliers. Later agreements between an importer and other suppliers therefore eroded the value of earlier concessions. This led to complex, acrimonious, and frustrating patterns of bargaining.³⁴

Despite these difficulties, reciprocity remains a valuable strategy for decentralized enforcement of cooperative agreements. Players who are aware of the problems of echo effects, bargaining deadlocks, and issue interdependence can compensate for these pitfalls. Axelrod observes that a better strategy than Tit-for-Tat "might be to return only nine-tenths of a tit for a tat."³⁵ The Tokyo Round dealt with the deadlock problem by beginning negotiations not on the basis of current tariff rates, but rather on the basis of a formula for hypothetical large across-the-board tariff cuts, with provisions for withdrawing offers on sensitive products, or if adequate compensation was not received. The problem of issue interdependence was dealt with in the trade area through multilateralization of tariff negotiations and adoption of unconditional most-favored-nation treatment.

These difficulties in applying reciprocity, and the responses of players to them, illustrate the significance of the institutions within which reciprocity is practiced. As noted above, multilateral trade negotiations are a case in point. In the military-security area, reciprocity has also been institutionalized. For example, stationing of American troops in Europe is linked to purchases of American military equipment by European governments. NATO as an institution has helped member governments achieve a variety of such reciprocal arrangements.

The debt negotiations discussed by Lipson also illustrate how reciprocity can be institutionalized in an N-person game. First, the major actors are identified, and bilateral negotiations take place between them or their agents. The I.M.F. and committees of banks negotiate with debtor countries. At a second stage, smaller banks are given the oppor-

³³ Evans, *The Kennedy Round in American Trade Policy: The Twilight of the GATT?* (Cambridge: Harvard University Press, 1971), 31-32.

³⁴ See Robert O. Keohane, "Reciprocity in International Relations," *International Organization* 40 (Winter 1986).

³⁵ Axelrod (fn. 6), 138.

tunity to adhere to these bargains, but not to influence their terms. At this stage, emphasis is placed on reciprocity at a different level: although the smaller actors have the potential to act as free-riders, efforts are made to ensure that they have incentives not to do so for fear that they may suffer in a larger game. Small banks face the threat of being excluded from crucial relationships with big banks, and from future lending consortia, if they fail to provide funds for rescheduling loans. As in the other cases described above, strategies of reciprocity for debt rescheduling are adapted creatively to avoid the problems of issue-interdependence that arise when there are many actors.

III. CONCLUSION

A. THE IMPORTANCE OF PERCEPTION

The contributors to *Cooperation under Anarchy* did not specifically set out to explore the role of perception in decision making, but the importance of perception has kept asserting itself. The significance of perception, including beliefs and cognition, will come as no surprise to students of international politics.³⁶ Yet it is worth pointing out once again that decision making in ambiguous settings is heavily influenced by the ways in which the actors think about their problem.

While this point has been made in laboratory studies many times,³⁷ there is an important twist in international politics that does not get sufficient attention from the psychologists who study decision making in the laboratory. Leaders of one state live far away from the leaders of other states. They are far away not only in space, but also in their cognitive framework: their tacit assumptions differ about what is important, what needs to be done, and who bears the responsibility for change. Put simply, those acting on behalf of states often do not appreciate how their own actions will affect others and how they will be interpreted by others. As Van Evera concludes from his study of World War I, preventing that war would have required dispelling extensive misperceptions that were prevalent in Europe before 1914.

Other striking examples of the importance of perception also come from the security area. For example, Downs, Rocke, and Siverson have found that even when nations in arms races built defensive rather than

³⁶ Jervis (fn. 31).

³⁷ For example, Amos Tversky and Daniel Kahneman, "Judgment under Uncertainty: Heuristics and Biases," *Science* 185 (September 1974), 1124-31; Richard Nisbet and Lee Ross, *Human Inference: Strategies and Shortcomings of Social Judgment* (Englewood Cliffs, NJ: Prentice-Hall, 1980).

offensive weapons, it was usually done not to defuse the arms race, but simply because they believed that such weapons offered the greatest amount of security per dollar. Even more to the point is that many arms races were started or accelerated without serious appreciation of the consequences. For example, when the Soviet leaders deliberately exaggerated their bomber strength in 1955 and their ICBM capabilities several years later, they did so for short-term political advantages; there is no evidence that they fully appreciated the long-term consequences that would follow when the United States geared up to take the threat seriously. In general, Downs, Rocke, and Siverson find that arms races are not often perceived as the result of actions chosen by others. In the events leading to the outbreak of war, national leaders may completely misunderstand the consequences of their acts. Van Evera notes, for example, that in 1914 the Russian government did not realize that Russia's mobilization would lead directly to Germany's mobilization, and to war. Another example of the impact of biased interpretations of events is provided by Jervis in his discussion of the decay of great-power concerts, which were undermined by divergent views of which side had made greater concessions to maintain cooperation.

While security issues provide the most dramatic examples, governments may be no better at understanding how their actions in the realm of political economy will be seen by others. Conybeare's study shows that trade wars have sometimes begun when states held mistaken beliefs that other countries would be reluctant to raise tariffs on imported food in retaliation for new tariffs placed on their exported manufactured goods. Trade wars have begun when states had exaggerated expectations about the tolerance of others for attempts at minor exploitation in widely accepted terms of trade.

B. GROPING TOWARD NEW INSTITUTIONS AND NORMS

Our project began with a set of hypotheses about how specific features of an international setting would affect the chances for the development of cooperation. Factors included were mutuality of interests, the shadow of the future, and the number of actors. These hypotheses have been supported by a broad set of cases that began in the 14th century, and covered trade disputes, monetary policy, and debt rescheduling as well as arms races, the outbreak of war, and diplomatic concerts. The three factors did, in fact, help to account for both cooperation and conflict.

We also discovered something else: over and over again we observed that the actors were not satisfied with simply selecting strategies based upon the situation in which they found themselves. In many cases we

saw deliberate efforts to change the very structure of the situation by changing the context in which each of them would be acting. Decision makers themselves perceived (more or less consciously) that some aspects of the situations they faced tended to make cooperation difficult. So they worked to alter these background conditions. Among the problems they encountered were the following:

1. how to provide incentives for cooperation so that cooperation would be rewarded over the long run, and defection punished;
2. how to monitor behavior so that cooperators and defectors could be identified;
3. how to focus rewards on cooperators and retaliation on defectors;
4. how to link issues with one another in productive rather than self-defeating ways and, more generally, how to play multilevel games without tripping over their own strategies.

A fundamental strategic concept in attaining these objectives is that of reciprocity. Cooperation in world politics seems to be attained best not by providing benefits unilaterally to others, but by conditional cooperation. Yet reciprocity encounters many problems in practice. As Axelrod has demonstrated, and as Van Evera's discussion of 1914 illustrates, payoff structures in the strategic setting may be so malign that Tit-for-Tat cannot work. Reciprocity requires the ability to recognize and retaliate against a defection. And retaliation can spread acrimoniously.

Actors in world politics seek to deal with problems of reciprocity in part through the exercise of power. Powerful actors structure relationships so that countries committed to a given order can deal effectively with those that have lower levels of commitment. This is done by establishing hierarchies, as one would expect from Herbert Simon's assertion that complex systems will be hierarchic in character.³⁸ In the present symposium, the construction of hierarchy for the sake of cooperation is best illustrated by Lipson's discussion of inter-bank networks to facilitate rescheduling of Third World debts; but it is also evident in Jervis's discussion of great-power concerts.

Another way to facilitate cooperation is to establish international regimes. Regimes can be defined as "sets of implicit or explicit principles, norms, rules, and decision-making procedures around which actors' expectations converge in a given area of international relations."³⁹ International regimes have been extensive in the post-1945 international

³⁸ Simon, *The Sciences of the Artificial* (Cambridge: MIT Press, 2d ed. 1982), chap. 4, "The Architecture of Complexity," p. 99.

³⁹ Krasner (fn. 8), 3.

political economy, as illustrated by the international trade regime (centered on the GATT) and the international monetary regime (including the I.M.F. as well as other organizations and networks).⁴⁰ Since the use of power can facilitate the construction of regimes, this approach should be seen as complementary to, rather than in contradiction with, an emphasis on hierarchical authority. Regimes do not enforce rules in a hierarchical sense, but they do change patterns of transaction costs and provide information to participants, so that uncertainty is reduced. Jervis argues that the Concert of Europe helped to facilitate cooperation by making it easier for governments to understand one another. Lipson shows how, in the regime for debt rescheduling, the control of information is used to facilitate cooperation on terms favored by the big banks. He also indicates that one weapon in the hands of those banks is their ability to structure transaction costs: the costs of negotiations involving major money-center banks are reduced while the costs of coordinating resistance by small banks are not. Conybeare's analysis implies that if England and the Hanseatic League had been able to form an international trade regime, they might have been able to make mutually advantageous bargains and to discipline some of their more unruly constituents.

International regimes do not substitute for reciprocity; rather, they reinforce and institutionalize it. Regimes incorporating the norm of reciprocity delegitimize defection and thereby make it more costly. Insofar as they specify precisely what reciprocity means in the relevant issue-area, they make it easier to establish a reputation for practicing reciprocity consistently. Such reputations may become important assets, precisely because others will be more willing to make agreements with governments that can be expected to respond to cooperation with cooperation. Of course, compliance is difficult to assure; and international regimes almost never have the power to enforce rules. Nevertheless, since governments with good reputations can more easily make agreements than governments with bad ones, international regimes can help to facilitate cooperation by making it both easier and more desirable to acquire a good reputation.⁴¹

International regimes may also help to develop new norms, as Ruggie has argued.⁴² Yet few such examples are evident in the cases discussed in this volume. The great-power concerts discussed by Jervis embodied new norms, but these did not last long; and the new norms of the 1930s

⁴⁰ Keohane (fn. 8), chaps. 8-9.

⁴¹ *Ibid.*, esp. chaps. 5-7.

⁴² Ruggie (fn. 15).

monetary system described by Oye were largely uncooperative and connected with the breakdown rather than the institutionalization of a regime. Major banks today are trying mightily to strengthen norms of repayment (for debtors) and of relending (for banks), but it is not at all clear that this will be successful. Better examples of creating norms may be provided by the evolution of thinking on chemical and biological warfare, and by the development, under GATT, of norms of non-discrimination—which are now, as we have seen, under pressure. Evidently, it is difficult to develop new norms, and they often decay in reaction to conspicuous violations.

Establishing hierarchies, setting up international regimes, and attempting to gain acceptance for new norms are all attempts to change the context within which actors operate by changing the very structure of their interaction. It is important to notice that these efforts have usually not been examples of forward-looking rationality. Rather, they have been experimental, trial-and-error efforts to improve the current situation based upon recent experience. Like other forms of trial-and-error experimentation, they have not always worked. Indeed, it is instructive to enumerate the variety of ways in which such experiments can fail.

1. The most important source of failure is that efforts to restructure the relationships may never get off the ground. As Downs, Locke, and Siverson note, there was an active peace movement in the years before 1914, and World War I was preceded by a series of conferences designed to secure arms control and strengthen international law; but these efforts did not significantly affect the nature of world politics. Similarly, the shakiness of monetary arrangements in the 1920s was perceived by many of the participants, but conferences to deal with these weaknesses, such as that at Genoa in 1922, failed to cope with them effectively. The great-power concerts discussed by Jervis seemed to get somewhat farther, but were never sufficiently institutionalized to have much prospect of longevity.
2. Some agreements are instituted, but turn out to be self-contradictory. We have noted that sequential bilateral negotiations under conditional most-favored-nation treatment may lead to a problem of infinite regress: each bargain tends to require the renegotiation of many others. Bilateral arms control agreements, whose restraints could encourage third parties to increase their armaments in order to catch up with the major powers, face a similar difficulty.
3. Even successful arrangements are subject to decay. Decay can result

from actors' attempts to find loopholes in established rules. The very success of GATT in reducing tariff rates contributed to an expansion of nontariff barriers; and efforts to evade those barriers led to their progressive extension and tightening.⁴³ Likewise, successful cooperation in the area of security may lead governments to believe that their partners' cooperation is not based on reciprocity but is unconditional. Insofar as this belief is incorrect, discord may ensue.

4. In some cases, changes that have nothing to do with the arrangements make them obsolete. Thus the international debt regime in place before the crisis of August 1982 was manifestly ill-equipped to handle a situation in which most Third World debts had to be rescheduled. In this instance, the old regime was adapted to meet new needs. The Depression of the 1930s made the monetary orthodoxy of the gold exchange standard obsolete. Indeed, Oye argues that the cooperative international monetary arrangements of the 1920s hindered attempts at monetary cooperation during the 1930s. The collapse of the old regime was a necessary condition for creation of a new one.

Eventually, any institution is likely to become obsolete. The question is under what conditions international institutions—broadly defined as “recognized patterns of practice around which expectations converge”⁴⁴—facilitate significant amounts of cooperation for a period of time. Clearly, such institutions can change the incentives for countries affected by them, and can in turn affect the strategic choices governments make in their own self-interest.

This interaction between incentives and institutions suggests the importance of linking the upward-looking theory of strategy with the downward-looking theory of regimes. The strategic approach is upward-looking in that it examines what individual actors will choose to do, and derives consequences for the entire system based on these choices. Most of the analysis in this volume has followed this upward-looking approach. On the other hand, much regime analysis has been downward-looking in that it examines the implications, for actors, of the way the entire system is organized. Some recent work has attempted to combine

⁴³ Vinod Aggarwal, “The Unraveling of the Multi-Fiber Arrangement, 1981: An Examination of Regime Change,” *International Organization* 37 (Autumn 1983), 617-46; David B. Yoffie, *Power and Protectionism: Strategies of the Newly Industrializing Countries* (New York: Columbia University Press, 1983).

⁴⁴ Oran R. Young, “Regime Dynamics: The Rise and Fall of International Regimes,” *International Organization* 36 (Spring 1982), 277-98; reprinted in Krasner (fn. 8), 93-114.

these two approaches,⁴⁵ but it has not yet been done in either a formally rigorous or an empirically comprehensive way.

The experimental groping by policy makers does not necessarily lead to stronger and ever more complex ways of achieving cooperation. The process proceeds by fits and starts. The success of each step is uncertain, and there is always danger that prior achievements will come unstuck. New experiments are often tried only under obvious pressure of events (as in debt rescheduling). And they are often dependent upon the active leadership of a few individuals or states who feel a serious need for change and who have the greatest resources.

The essays in this collection show that we are beginning to understand the structural conditions that affect strategic choices leading to cooperation or discord. These factors are mutuality of interest, the shadow of the future, and the number of actors. Over a wide range of historical cases, these three dimensions of situations do help account for the emergence, or nonemergence, of cooperation under anarchy.

But in the course of this collective research we have also found that states are often dissatisfied with the structure of their own environment. We have seen that governments have often tried to transform the structures within which they operate so as to make it possible for the countries involved to work together productively. Some of these experiments have been successful, others have been stillborn, and still others have collapsed before fully realizing the dreams of their founders. We understand the functions performed by international regimes, and how they affect strategies pursued by governments, better than we did a number of years ago. What we need now are theories that account for (1) when experiments to restructure the international environment are tried, and (2) whether a particular experiment is likely to succeed. Even within a world of independent states that are jealously guarding their sovereignty and protecting their power, room exists for new and better arrangements to achieve mutually satisfactory outcomes, in terms both of economic welfare and military security.

This does not mean that all endeavors to promote international cooperation will yield good results. Cooperation can be designed to help a few at the expense of the rest; and it can accentuate as well as alleviate injustice in an imperfect world. Yet the consequences of failure to co-

⁴⁵ In *After Hegemony* (fn. 8), Robert Keohane has sought to show how game theory (which is "upward-looking") can be combined fruitfully with the "downward-looking" theories of public goods and market failure to develop a functional theory of international regimes. But he has not formalized his theory, and has applied it only to the post-World War II international political economy.

operate—from warfare to the intensification of depressions—make us believe that more cooperation is often better than less. If governments are prepared to grope their way toward a better-coordinated future, scholars should be prepared to study the process. And, in a world where states have often been dissatisfied with international anarchy, scholars should be prepared to advance the learning process—so that despite the reality of anarchy, beneficial forms of international cooperation can be promoted.