ARMS RACES AND COOPERATION

By GEORGE W. DOWNS, DAVID M. ROCKE,
and RANDOLPH M. SIVERSON*

I. INTRODUCTION

A GOVERNMENT interested in reducing the level of arms com-
petition with a rival can attempt to realize this goal in a variety
of ways. It can choose to shift military expenditures from offensive to
defensive weapons in the hope that this will solve the "security dilemma"
by increasing its security without decreasing that of its opponent. It can
attempt to convince its rival of the futility of increased arms expenditures
by announcing that it will henceforth spend neither more nor less on
defense than the rival. Or it can propose that the two states initiate
formal arms talks. Each of these actions represents a different type of
strategy designed to promote cooperation. Shifting expenditures to pri-
marily defensive weapons is one type of *unilateral* strategy that has been
discussed at length by Jervis.[1] Tying defense expenditures to those of
the rival state is an example of the *tacit bargaining* strategy of reciprocity
that plays a critical role in Schelling and in Axelrod's recent work on
the Prisoners' Dilemma.[2] Formal arms talks are a variety of the strategy
of *negotiation*.[3]

In this paper we explore the advantages and disadvantages of each
of these strategies for reducing the intensity of arms races that are
motivated by different patterns of preferences and complicated by dif-
ferent sources of uncertainty. The essay contains a substantial amount
of both formal and historical analysis. To rely on formal analysis alone
would tend to obscure the fact that arms races have taken place under
a far broader set of conditions than might be predicted on the basis of

¹ Robert Jervis, *Perception and Misperception in International Politics* (Princeton: Princeton
University Press, 1976); and, more notably, Jervis, "Cooperation under the Security Di-
lemma," *World Politics* 30 (January 1978), 167-214.
² Robert Axelrod, *The Evolution of Cooperation* (New York: Basic Books, 1984).
³ Howard Raiffa, *The Art and Science of Negotiation* (Cambridge: Harvard University
Press, 1982).

a narrow game-theoretic representation, and that they can be affected by factors that are not conventionally dealt with in formal analysis. For example, when states employ "bargaining chip" strategies or attempt to convince a rival that they are stronger than they really are in order to gain policy leverage in other areas (as the Soviet Union did in the late 1950s), an arms race may develop or accelerate even though the structure of the payoffs suggests that cooperation should result. Moreover, we will see that it is possible to propose a strategy that can successfully promote cooperation in a world of perfect information and control, but fail miserably in the uncertain world in which states must interact.

An approach that attempts to evaluate alternative cooperative strategies solely by examining the historical record of the 19th and 20th centuries faces still greater problems. This is not, as one might assume, simply because there are relatively few instances in which arms races have been peacefully reversed. Although their number is not large enough to isolate the impact of an individual strategy in cases where several were pursued simultaneously, they are more numerous than many pessimists might imagine.[4] The chief difficulty lies instead in the disconcerting fact that the majority of arms races that have ended in cooperation have done so not because one side adopted a particular cooperative strategy, but because the basic character of the race was altered by events that were not directly connected with it.

Such events, which collectively might be called the "tilt of the board," spring largely from economic circumstances and changes in the larger political order outside the relations between the two states engaged in the arms race. From a historical point of view, the most important determinant of cooperation that can be classified under this heading is the activity of a third power. Consider the series of Anglo-French naval races that took place during the last half of the 19th century. The race of 1852-1853 was "peacefully" resolved when the two nations joined

---

[4] Both Samuel Huntington, in "Arms Races: Prerequisites and Results" *Public Policy* 8 (1958), 41-86, and Paul Kennedy, in *Strategy and Diplomacy* (Ayelsbury, England: Fontana, 1984), 163-78, offer examples of arms races that did not result in war. Huntington's list, which is the more comprehensive of the two, follows:

1. France vs. England	1840-1866
2. France vs. Germany	1874-1894
3. England vs. France and Russia	1884-1904
4. Argentina vs. Chile	1890-1902
5. England vs. Germany	1898-1912
6. England vs. United States	1916-1930
7. Japan vs. United States	1916-1922
8. United States vs. Soviet Union	1946-

Kennedy, who concentrates his attention upon only the major powers, recognizes all of the above except Argentina and Chile between 1890 and 1902; there are minor differences in some of the dates.

forces to fight the Russians in the Crimean War; that of 1859-1861 gradually dissipated as Napoleon III turned his attention and limited resources to coping with Mexico and the rising power of Prussia; and that of 1884-1904 ended in the face of increased German power and aggressiveness.[5] Although these races came to a halt without the rivals going to war (at least with each other), they hardly provide ideal cases for making inferences about the effectiveness of alternative strategies to induce cooperation. Fortunately, while the relatively minor role played by cooperative strategies in resolving past arms races limits our ability to evaluate their merits directly, that does not mean that historical analysis has nothing to contribute. As we shall see, it can substantially enrich the formal analysis of arms races by identifying variables that might otherwise be omitted, and by prompting speculation about which of the wide variety of "games" that can theoretically lead to an arms race are actually being played.

II. Arms Races, Perfect Information, and Cooperation

To begin to understand the merits of the three strategies and the important variations in arms races with which they may have to cope, it is useful to consider the least complicated abstraction of two utility-maximizing antagonists with perfect information about each other's preferences and behavior. Let us treat an arms race as a sequence of plays of a 2 x 2 game with the standard payoff matrix shown in Figure 1.[6] To understand the conditions on the payoffs necessary for an arms race to exist, consider the payoffs to one of the participants.[7] In this setting of full information, an arms race—defined in this context as continual mutual defection—cannot exist among rational opponents un-

[5] Richard Cobden, *The Political Writings of Richard Cobden* (London: William Ridgway, 1868); Kennedy (fn. 4); Arthur J. Marder, *The Anatomy of British Sea Power* (New York: Octagon Books, 1976).

[6] The key here is the preference structure; for our purposes, the other conditions for a 2 x 2 game do not necessarily have to hold. The sequential-play-of-the-game structure seems the most reasonable in spite of Wagner's argument that one should allow each player to act conditionally on knowledge of the other's play. See R. Harrison Wagner, "The Theory of Games and the Problem of International Cooperation," *American Political Science Review* 77 (June 1983), 330-46. This seems more like defining away the problem than solving it: if information moves quickly enough and control is fine enough so that one can respond to an opponent's move after one week, then one models this as a-game-a-week. The analysis in extensive form that is used by Wagner seems unnecessarily cumbersome.

[7] In order to reduce the number of cases that have to be discussed, the analysis here is carried on with respect to the payoffs of one of the participants rather than to those of the whole game. Once the analysis of the individual orderings has been made, the actual games can be constructed using combinations of preference patterns. Terms such as "Prisoners' Dilemma," which conventionally refer to symmetric games, are also used for the preference pattern of one of the participants.

FIGURE I

2 X 2 REPRESENTATION OF AN ARMS RACE

Country B

		Cooperate	Defect
	Cooperate	C,C	C,D
Country A	Defect	D,C	D,D

less DD > CD; otherwise the response to defection would be cooperation.[8] If this requirement is fulfilled for both participants, we say that the game is defection-stable in the sense that once either party defects, both will defect from that point on. A second requirement for an arms race between rational opponents is that CC > DC and CC > DD cannot both be true. If CC > DC, cooperation is preferred to unilateral defection, so there is no incentive to defect from mutual cooperation (a condition described as cooperation stability); if CC > DD, this stable outcome is preferred to mutual defection.[9]

The following nine orderings of the outcomes for each side fulfill the above conditions:[10]

1. DD > CC > CD > DC
2. DC > CC > DD > CD   (Prisoners' Dilemma)
3. DD > CC > DC > CD
4. DD > CD > CC > DC
5. DC > DD > CC > CD   (Deadlock)
6. DD > DC > CC > CD   (Deadlock-type)
7. DD > CD > DC > CC
8. DC > DD > CD > CC   (Deadlock-type)
9. DD > DC > CD > CC   (Deadlock-type)

[8] The game Chicken, which violates this condition, has been used by several authors for representing conflict situations. For an example, see Glen Snyder and Paul Diesing, *Conflict Among Nations: Bargaining, Decision-Making, and System Structure in International Crises* (Princeton: Princeton University Press, 1977). Chicken is highly unstable, however, and not a realistic model for an arms race as *continual* defection.

[9] We shall see later that some of these excluded games can lead to arms races when the assumption of perfect information is relaxed.

[10] The terminology for the various games is presented in Snyder and Diesing (fn. 8) and Jervis (fn. 1). Snyder and Diesing use a different selection of games since they are interested in international crises—events of short duration—rather than the more extended arena of arms races. Jervis focuses mostly on Stag Hunt, which is an important sub-case that we deal with in a subsequent section.

A. DEADLOCK

Four of these orderings are related to the game of Deadlock in that the two most preferred outcomes involve defection by the side whose preferences they reflect. Because Deadlock provides so little opportunity for the generation of a cooperative solution, and because such a single-minded preference for defection seems to make war inevitable, there has been a persistent tendency on the part of social scientists to imply that it is descriptive of no more than a handful of deviant cases (for instance, Hitler's Germany). Such an apocalyptic view of Deadlock may be naive for a number of reasons. Contrary to the theories of Richardson[11] and others, there is no compelling evidence that even substantial periods of continuing arms increases necessarily lead to war.[12] They may persist in an uncomfortable but generally peaceful, dynamic equilibrium or, more commonly, be brought to a peaceful conclusion by some exogenous shock. More important, Deadlock and Deadlock-type preferences may be characteristic of decision makers who share none of Hitler's imperialistic dreams. Consider, for example, a situation where internal constituencies press for military expenditures to increase the well-being of the arms industry. If this internal component carries sufficient weight in determining defense expenditures, the preferences of decision makers will be those of Deadlock: they will choose to defect regardless of the action of the rival state. In this case, their preferences are those of Deadlock—not because of an insatiable lust for war, but because they place a premium on a particular domestic constituency.

Nor is this the only situation in which Deadlock preferences may arise. If a government believes that its weapon systems are superior to those of its rival, or that the strategy that will guide their use is superior, it may prefer mutual defection to mutual cooperation. Whether it does or not depends (at least under our assumptions of rationality) on which option offers the greatest net benefit. There is some evidence to suggest that, before World War I, the Germans' belief in the superiority of their dreadnoughts and crews led to a perception that a matching shipbuilding effort on the part of the British was preferable to mutual cooperation. This belief was probably reinforced by another sort of utility calculation

[11] Lewis F. Richardson, *Arms and Insecurity* (Chicago: Quadrangle, 1960).

[12] Recent research by Michael Wallace draws a very close connection between arms races and war. See Wallace, "Arms Races and Escalation: Some New Evidence," *Journal of Conflict Resolution* 23 (March, 1979), 3-16, and "Armaments and Escalation" *International Studies Quarterly* 26 (March 1982), 37-56. Subsequent analysis casts very serious doubt on the relationship, however; see Henk W. Houweling and Jan G. Siccama, "The Arms Race—War Relationship: Why Serious Disputes Matter," *Arms Control* 2 (September 1981), 157-97.

that seems to have played a critical role in a number of arms races: that of relative strength. If state A has 10 tanks and state B has 50, the decision of A to acquire 5 tanks increases its total forces by 50 percent and increases its strength relative to B by 30 percent. Should B choose to match this increase, A still increases its relative strength by 27 percent. Thus mutual defection leaves A nearly as well off as unilateral defection, and possibly better off than mutual cooperation. The logic of this situation constituted the underpinnings of Tirpitz's German naval policy as it developed after 1899, and almost certainly played a role in defining the Deadlock-like preferences of the Germans in the 1930s and of the Soviets in the 1950s. Both began from a position that was so weak relative to their natural rivals that it could reasonably be expected to improve under mutual defection.

Finally, the presence of Deadlock or Deadlock-type preferences would help to explain the most striking findings associated with the evolution of cooperation in 19th- and 20th-century arms races: the relative scarcity of national strategies to end them, and the important role of exogenous events. What easier way to explain why nations so infrequently employed one of our three strategies to promote cooperation than to hypothesize that their preference was to arm regardless of what their rival did? What simpler way to explain why outside factors proved critical in ending an arms race than to suggest that the participants had no other incentive to do so? Although there are other explanations for both findings, their joint occurrence suggests the possibility that Deadlock is far more prevalent than is usually thought to be the case.

Let us now consider the ability of each of the three cooperative strategies to cope with a Deadlock-driven arms race. We will begin with the unilateral strategy, which increases the potential for cooperation without requiring any active participation or consensus on the part of the rival state. It does this by providing a state with the benefits of security that it would normally get from arms without provoking the rival to increase its armaments and therefore the intensity of the arms race.[13] A unilateral strategy possesses the virtue of leaving the initiating state better off (in the sense of gaining security) than it would be under a variety of other more formal cooperative arrangements, such as a bilateral freeze.

There are at least four types of unilateral strategies that can reduce the intensity of an arms race. The first (and the one that has received the most theoretical attention) involves a concentration on defensive as

[13] Jervis (fn. 1).

opposed to offensive weapons.[14] Recent proposals to increase the invul-
nerability of ICBMs to preemptive attack by hardening silos, employing
mobile launch platforms, or developing a system of silo-defense ABMs
instead of simply increasing the number and accuracy of ICBMs, have
all been justified at least partly on the grounds that they will make the
United States more secure without provoking a new round of arms
escalation.[15] Curiously, historical examples of defensive weapons being
acquired because of their ability to defuse an arms race seem to be quite
rare. States have often chosen to build defensive rather than offensive
weapons during an arms race, but they usually appear to have done so
because they believe that such weapons offered the greatest amount of
security for the money spent, and not because they thought that such
weapons would lead to an end of the arms race. Thus, the British adopted
a policy of building picket ships and fortifying the coast during their
naval race with France in 1859-1861 without giving any particular
thought to how the French would respond, or what impact it would
have on their arms race.[16] These moves simply seemed to provide greater
security than could be obtained through an exclusive reliance on ships
of the line. The same absence of any unilateral strategy logic charac-
terized Britain's defensive response to Germany's bomber build-up prior
to World War II. The British chose to build fighters, radar, and anti-
aircraft guns rather than bombers—not because they were afraid of
provoking the Germans to still greater bomber production, but because
they thought it was the optimal way to compete with them.[17]

   A second type of unilateral strategy that can be used to break out of
the security dilemma is the formation of defensive alliances. In the purest
case, these function in precisely the same fashion as the acquisition of
defensive weapons. The only difference is that, in this instance, defensive
weapons are being "purchased" through the formation of an alliance
instead of directly, thus permitting states to increase or maintain their
level of security without engaging in the budgetary expenditures that
weapons increases would entail. Just as in the case of a shift from
offensive to defensive weapons, however, there are many instances of
defensive alliances, but there have been few one could confidently assert
to have been undertaken as part of a strategy to solve the security
dilemma. When France and Russia entered into their alliance of 1894,
they were well aware that their defensive capabilities were significantly

[14] *Ibid.*
[15] Freeman Dyson, *Weapons and Hope* (New York: Harper & Row, 1984).
[16] Cobden (fn. 5), 367.
[17] Ian Colvin, *The Chamberlain Cabinet* (New York: Taplinger, 1971).

enhanced; but they did not believe that their alliance would leave Great Britain (or Germany) undisturbed—which indeed it did not. The Byzantine pattern of defensive alliances that characterized Bismarckian Europe appears to have been more successful in preventing sustained arms races; but again, it is not clear that this was the primary intent of the actors.[18]

A third type of unilateral strategy involves the creation of buffer states. Ideally, this strategy provides both sides with greater security simultaneously. The creation of a neutral and independent Belgium in 1832 is a good example, although here the motivation had less to do with defusing an ongoing arms race than with avoiding one.[19]

A fourth type of unilateral strategy is based on the acquisition and surrender of intelligence information. The more accurate the information the United States possesses about Soviet weapons development, and the more quickly it is obtained, the more costly it is for the Soviets to try to escalate the arms race covertly. Providing the enemy with intelligence information voluntarily in order to convince it that its worst fears are unjustified is rare, but it has occurred. In 1853, the French Minister of Marine invited members of Parliament to visit French shipyards to allay British suspicions of a massive naval build-up.[20]

Because Deadlock creates no incentive to cooperate, it may appear that the various unilateral strategies are irrelevant to the problem of arms race de-escalation in such an environment. Although this is true in the sense that, under Deadlock, states do not choose to pursue a unilateral strategy in order to promote cooperation, they can nonetheless choose to do so from other motives with similar effect. Suppose, for example, that the motivation that lies behind a state's present preference for defection is a desire for greater security. If it believes that the net benefit in added security to be gained from an increased expenditure on defensive arms is greater than it would be if it cooperated and

<hr />

[18] A.J.P. Taylor, *The Struggle for Mastery in Europe*, 1848-1914 (London: Oxford University Press, 1971).

[19] As British Foreign Minister Palmerston wrote after the London Conference:
France ought to feel the great advantages which this arrangment confers upon Her. By one stroke of the pen, the whole line of Belgick fortresses, so far as they constituted points of attack upon her territory, at once disappear; and this upon a frontier the nearest to her Capital, and the least protected by defenses of nature or of art. England voluntarily interdicts Herself ... while the Northern powers of Europe, of their own accord, close the door through which they would naturally approach the French frontier in the event of hostilities leading them there. France ought therefore to see in this Protocol the most signal and unequivocal proofs of the pacific spirit which animates the other Powers of Europe ...
Quoted in William Lingelbach, "Belgian Neutrality: Its Origin and Interpretation," *American Historical Review* 39 (October 1933), 60-61.

[20] Cobden (fn. 5), 259.

maintained the status quo, then it would adopt Deadlock-like preferences, but *behave* as if it had adopted a unilateral strategy. As we have noted above, this is what seems to have been going on in the case of British fortress and picket-ship construction. The behavior resembled the implementation of a unilateral strategy, but the motive was Deadlock: the British preferred building forts and ships to mutual cooperation with the French. It also seems to describe the logic of various members of the Reagan administration with regard to the antiballistic missile. That is, they appear inclined to proceed with an ABM system regardless of whether the Soviets increase the number and sophistication of their ICBMs, because they believe the United States would be better off with such a system than it would be under a nuclear freeze. In cases like these, the motivation is security rather than cooperation, but the outcome is the same. The successful implementation of a perimeter defense system in the case of the British, or an antiballistic missile system in the case of the United States, reduces the marginal benefit of shipbuilding by the French or the manufacture of ICBMs by the Soviets. If the French and the Soviets were also motivated primarily by security considerations rather than a desire to expand their ability to influence their rivals, these actions should slow the respective arms races.[21]

Tacit bargaining differs from a unilateral strategy in that the actions taken are made dependent on those of the other state; it differs from conventional negotiation in that the role of verbal interaction is relatively insignificant, and no attempt is made to create a formal agreement. Perhaps the closest parallel is the reinforcement schedule used in behavior modification to reward and punish a subject in such a way that he or she eventually "learns" a desired behavior. Of course, each state here is simultaneously both "subject" and "experimenter." Tacit bargaining is not relevant to Deadlock-driven arms races, since neither state has an incentive to cease defecting and to send a positive signal or, for that matter, to pay attention to such a signal if it were sent.

---

[21] The question of whether it will in fact slow the arms race is a complex one. One could argue that the other side will build fewer offensive weapons since the marginal benefit of each is reduced; one could also argue that more will be built since more will be required to obtain a given objective. Without further assumptions, either result is possible. Suppose that one side established defenses so that each weapon of the other side is 10% less effective than it was. This has two simultaneous effects. First, it raises by 10% the price of what might be called a weapon effectiveness unit; that is, the amount of a weapon required to accomplish a given objective. Second, it increases the number of weapons required per weapon effectiveness unit by 10%. If the price elasticity of demand for weapon effectiveness units is −1, then the 10% increase in price will result in a 10% decrease in the number of weapon effectiveness units purchased. However, since 10% more weapons are needed for each effectiveness unit, the number of weapons remains the same. If this elasticity is greater than −1 (say −0.5), the number of weapons built will increase; whereas, if the elasticity is less than −1 (say −1.5), the number of weapons built will decrease.

Negotiation, like tacit bargaining, may seem to hold little promise for reducing the intensity of a Deadlock-motivated arms race. Why negotiate if mutual defection is preferable to mutual cooperation? The answer lies in the fact that, while both states may prefer defection to cooperation within the confines of a particular game (i.e., the missile game, the dreadnought game, etc.), negotiation offers the possibility of linking that game to other issues (e.g., access to raw materials, recognition of territorial claims) in such a way that the net marginal benefit that each gains from cooperating within this larger game is greater than that of defection. Expressed more formally, the presence of additional issues transforms a basically zero-sum distributive game with no zone of agreement into a non-zero-sum integrative game in which both sides can benefit simultaneously.[22]

The trick is that, whatever these goods or policies are, they cannot have been freely exchanged prior to the negotiation. In that case, they would already have been exchanged for each other directly at an optimal level and there would be nothing to gain by introducing defense issues. The goods and policies must be linked to defense questions in such a way that an optimal level of exchange cannot take place in the absence of a negotiated arms agreement. In most arms races, they are not difficult to find. The U.S. and other states have frequently restricted the trade of raw materials, petroleum products, and various finished goods to countries with which they are engaged in an arms race. Any of these might plausibly be the basis for negotiating the end of a deadlocked arms race.

The role that sub-optimally exchanged goods and policies can play in ending a deadlocked arms race suggests that a national leader who has placed numerous economic and social restrictions on the state's interactions with a rival may be more able to negotiate the end of a deadlocked arms race than a leader who has imposed no such restrictions. This may help to explain why leaders with a reputation for great mutual hostility are often able to negotiate a settlement where their more moderate colleagues have failed. It is not simply because, as it is usually argued, they are both in a position to coopt their conservative followers as well as their more liberal opponents, but because they have so thoroughly linked other issues to the arms race that they have more to offer each other in return for cooperation.

B. PRISONERS' DILEMMA

Each of the cooperative strategies can be effective in coping with an arms race inspired by a Prisoners' Dilemma. Moreover, because mutual

[22] Raiffa (fn. 3), 131 ff.

cooperation in a Prisoners' Dilemma is more attractive than mutual defection, there will be an incentive to invoke these strategies. There are two problems with unilateral strategies in the context of the Prisoners' Dilemma with perfect information. One lies in the implicit assumption that the motivation of the rival is primarily that of security. If this assumption is correct, the rival state will not be motivated to respond to the initiating state's acquisition of defensive weapons, a defensive alliance, and so forth, because it will feel no more threatened than it did prior to such an action. If, on the other hand, the rival is motivated by a desire to exert influence over the other's policies, it will respond to a build-up of defensive weapons or a defensive alliance in roughly the same way (in terms of effort if not technology) as it would to a build-up of offensive weapons or the creation of an aggressive alliance.

The other problem with unilateral strategies is that a cost-effective alternative to the development of offensive weapons is assumed to exist. Frequently, however, this is simply unavailable, or decision makers do not think it is available. Before the development of the submarine there was no defensive alternative to the acquisition of capital ships, just as there has been no real alternative to the ICBM. Similarly—although one could argue that the problem was perceptual—on the eve of World War I national leaders saw no real alternative to responding to troop build-ups by other countries except with build-ups of their own.

The virtues of alternative tacit bargaining strategies in contexts of Prisoners' Dilemma have recently been explored in a most interesting way by Axelrod.[23] He found that a policy of reciprocity—or Tit-for-Tat—that begins with cooperation and thereafter prescribes a response that reflects the rival's behavior achieved the best results when competing with a host of other strategies in an iterative Prisoners' Dilemma computer simulation. This strategy is effective, not because it links the game to other issues, but because it forges a link between current behavior and future iterations of the same game. The "defect" strategy is clearly the one for both players engaging in a single-play Prisoners' Dilemma game, since it maximizes the individual benefit of each regardless of what the other chooses. That is often not the case, however, in a mutiple-play version of the same game where the number of plays is large and unknown. Here the advantages of a defection in any single game can be outweighed by the expectation that such a decision will cause the opponent to respond by defecting in subsequent plays. This can be costly because the gain to be obtained from a single defection may be quite modest when compared to the difference between mutual cooperation and mutual defection if the latter is multiplied by a large number that

[23] Axelrod (fn. 2).

represents the times the game is likely to be played in the future. Axelrod points out that this explains why firms only infrequently violate their contracts even in contexts where legal recourse is impractical. The gain from any single violation is small compared to the gain to be realized through subsequent dealings with each other.

The prospect of iteration plays a central role in arms decisions because there is always the possibility that any rate of arms increase or technological innovation will inspire a new race or intensify the existing one. Everything else being equal, the larger the number of expected iterations in a Prisoners' Dilemma arms game, the more likely it will be that the short-term benefits of an arms lead will be judged to be less than the long-term costs of a protracted arms race. A government interested in promoting cooperation can take advantage of this fact by convincing its rival that it will respond aggressively to any arms escalation by increasing its own armaments, and that it has the resources to do so indefinitely. That is very close to the official policy adopted by Britain at a number of junctures during the 19th century, perhaps most explicitly in the "two-power standard" by which the British government announced that it would maintain a navy that was at least as large as the combined forces of the next two most powerful fleets. The power of iteration to promote cooperation can also be seen in the tendency of arms agreements to focus on weapons that would make continued iterations especially costly (e.g., dreadnoughts rather than destroyers) and on new, unproven weapon systems that yield only modest short-run benefits for the "lead" state and promise vastly increased costs in the future (e.g., ABMs, space weapons).

An interesting but widely neglected aspect of increased intelligence capabilities is that they not only reduce the short-term advantages of arms increases and innovations by permitting one state the opportunity to catch up to the other before the difference in capabilities is strategically significant, but that they further increase iterativeness by providing an opportunity for many timely responses. As we shall see, this can be quite beneficial in the execution of tacit bargaining strategies.

Of course, in security matters it is not always possible for decision makers to convince themselves, much less an opponent, that the long-term benefits of cooperation outweigh the short-term benefits of defection. As Van Evera makes clear, the cult of the offensive created a situation prior to World War I in which each government believed that a slight advantage in forces or weaponry could quickly be translated into a military victory.[24] Under these circumstances, the long-run advantages of cooperation (which Axelrod terms "the shadow of the fu-

[24] Stephen Van Evera, "Why Cooperation Failed in 1914," pp. 80-117 of this collection.

ture") pale before the benefits of victory or the cost of defeat, and the threat of future non-cooperation becomes irrelevant. Obviously this is the same argument that is used to justify an absence of cooperation in the nuclear arms race. Not surprisingly, perhaps, there seem to have been few if any arms races in the last 200 years where the rhetoric of government officials on one or both sides has not been filled with warnings that everything will be lost if there is any delay in matching the arms growth of the rival. For example, in the 19th century the cost of a future arms race cast only a small shadow in Great Britain, when news of the first French ironclad raised the specter of imminent invasion across an "iron bridge" spanning the English Channel.[25]

As in the case of Deadlock, the key to the capacity of negotiation to cope with arms races inspired by Prisoners' Dilemma lies in its ability to transform the game (or at least reduce the relative benefits of defection) by creating linkages to other games and issues. Sometimes, when the advantage of unilateral defection over mutual cooperation (DC − CC) is small and the advantage of mutual cooperation over mutual defection (CC − DD) is large, this linkage may be modest and implicit. If state A violates the agreement, state B will go back to defecting (just as in the case with tacit bargaining) or refuse to negotiate on other issues where there is a mutual advantage that might be realized. At other times the linkage is made more formally. In either case, there must be a capability to withdraw the issue or game to which the linkage is made if a defection takes place; otherwise there is no incentive to continue cooperating.

While one is hard-pressed to find historical examples of arms races that have been partially de-escalated through the use of a unilateral strategy or tacit bargaining, there appear to be three instances in which negotiation has played an important role in settings in which the participants can plausibly be argued to have Prisoners' Dilemma preferences. These are the Washington Naval Treaty of 1922 and SALT I and II. The Washington Naval Treaty of 1922 put a limit on the battleship-building efforts of what were then the world's three great naval powers: Great Britain, the United States, and Japan. To assess the contribution of negotiation, we need to know how complicated the agreement was that produced the compromise of the capital-ship tonnage ratio of 5:5:3, and to what extent issues over and above that of battleship construction were necessary to bring it about. In answer to the first question, the agreement involved a number of Pacific security and trade issues un-

[25] Cobden (fn. 5).

related to battleship construction.[26] Great Britain and the United States received guarantees of an open-door policy regarding trade in China and the Pacific, and the return of Shantung province and key transportation systems to the Chinese. They also received guarantees that, with certain specified exceptions, Japan would not increase fortifications on the Pacific islands taken from Germany during World War I. The latter provisions were especially important to the U.S., which found itself controlling areas such as the Philippines and Guam that would have been virtually indefensible if the Pacific were dotted with Japanese naval bases. Both sets of guarantees were also important to the British since they could now avoid having to keep a large navy in the Pacific. In addition, the Washington Treaty provided the British with an excuse not to extend the possibly entangling 1902 Anglo-Japanese alliance that was about to expire. For its part, Japan received the *de jure* right to administer the possessions north of the equator that it had taken from Germany, and was able to head off the possibility of a mutual assistance pact in the Pacific between Great Britain and the United States.

It is plausible that no capital-ship treaty would have been signed in the absence of these other agreements. Great Britain could not have afforded to reduce its presence in the Pacific without settlement of the China question. The United States, while probably the power least wedded to battleships as the core of its navy, was in the process of building more such ships (see the 1916 Defense Bill and the 1919 Defense Appropriations Bill); it might have increased their number still further if it had been necessary to defend its Pacific possessions and free trade in the East. The Japanese, who had only recently reconfirmed their commitment to a policy of two squadrons of eight battleships each, would almost certainly have built more ships sooner if faced with an Anglo-American Defense Pact (a likely outcome in the absence of the Washington Treaty) and continued U.S. insistence that they abandon the possessions taken from Germany.[27] The comments of Secretary of State Stimson to Senator Borah in a letter 10 years later attest to the importance of the multidimensional character of the final settlement:

It must be remembered also that this treaty was one of several treaties and agreements entered into at the Washington Conference by the various

[26] Harold Sprout and Margaret Sprout, *Toward a New Order of Sea Power: American Naval Policy and the World Scene, 1918-1922*, rev. ed. (Princeton: Princeton University Press, 1943; reprinted by Greenwood Press, 1969); Robert A. Hoover, *Arms Control: Interwar Naval Limitation Agreements* (Denver: Monograph Series in World Affairs, University of Denver, Vol. 17, Book 3, 1980).

[27] Quincy Wright, *The Existing Legal Situation as It Relates to the Conflict in the Far East* (New York: Institute of Pacific Relations, 1939).

powers concerned, all of which were interrelated and interdependent. No one of the treaties can be disregarded without disturbing the general understanding and equilibrium which were intended to be accomplished and effected by the group of agreements arrived at their entirety. The Washington Conference was essentially a disarmament conference aimed at the possibility of peace in the world, not only through the cessation of competition in naval armament, but also by the solution of other disturbing problems which threatened the peace of the world, particularly in the Far East.[28]

Thus, in the case of the Washington Treaty, additional dimensions had to be added to the capital-ship problem for an agreement to emerge; the result was complicated enough to have required negotiation.

To a somewhat lesser extent, the SALT I agreement was characterized by similar complexities. It contained provisions that allowed each state significant differences in the levels and "mixes" of its strategic arsenal. Even the most basic trade-off that permitted the Soviet Union an advantage in the area of throw weight, and the U.S. a greater number of missiles, would have been difficult to attain through tacit bargaining. Salt II, with the infusion of cruise missiles, Soviet bombers, and the issue of forward bases in Western Europe, was even more complicated. Talbott, for example, recounts the trade in which the maximum level of 1200 MIRVed warheads preferred by the U.S. was accepted by the U.S.S.R. in exchange for an American acceptance of the Soviet preference for a limit of 2,250 strategic launchers.[29]

The importance of being able to link different weapon systems and different dimensions of the same system, and to introduce non-weapon dimensions such as territory and trade, is also demonstrated by the fate of disarmament talks in which the negotiating parties did not have these options at their disposal. In the first Hague conference of 1899, the agenda was constructed in a way that carefully separated weapons limitation talks from discussions of every other issue; the representatives were primarily military experts who had neither the authority nor the expertise to introduce political, economic, or territorial issues. While it would be foolish to argue that the conference would have been a success if this situation had been different, it is not unreasonable to say that the narrowness of the mandate under which the representatives operated helped to create a problem of structural control that made the achievement of any real progress nearly impossible.[30]

[28] Ibid., 102.
[29] Strobe Talbott, Endgame: The Inside Story of SALT II (New York: Harper & Row, 1979).
[30] Calvin DeArmond Davis, The United States and the First Hague Peace Conference (Ithaca, NY: Cornell University Press, 1962).

III. Sources of Uncertainty:
Problems of Information and Control

The decision environments in which arms races actually take place pose a number of information and control problems which increase the range of games that can lead to arms races and affect the ability of any given strategy to produce cooperation. They are somewhat analogous to the market imperfections that play such an important role in public finance, in the sense that they both complicate the decision problem—in this case, obtaining cooperation in an arms race rather than maximizing marginal social benefit—and reduce the applicability of purely deductive analysis. Moreover, just as the importance of externalities, natural monopolies, or consumer ignorance is determined less by logic than through the observation of real markets, so the discovery of these complicating characteristics is grounded in the evolution of actual arms races. We will begin by enumerating several problems of information and control and considering how they can function to expand the range of games that can lead to arms races. Then we will show how they can cause the conduct of other games to vary from how they would be played under the assumption of perfect information.

A.   STRATEGIC MISREPRESENTATION

Strategic misrepresentation takes place in a bargaining situation when, in order to obtain an advantage, one party pretends to have preferences that it does not. In the context of arms negotiation, the classic situation is one in which a state pretends to want a weapon system more than it actually does, in an attempt to extort concessions from its rival for halting its development. Unfortunately such bluffs are sometimes called, in which case the desire to maintain future credibility may require that the threat to build the system be carried out. *This means that under sufficiently unlucky circumstances almost any game can result in an arms race.* Schelling's classic example of the ABM provides one illustration; another may be found in connection with the MX missile. One can argue that a majority of Congress believes that the nonmobile MX is unnecessary or undesirable even without Soviet restraint. That is, a plausible pattern for congressional preferences in the absence of the missile's role as a bargaining chip to encourage Soviet concessions might be CC > CD > DC > DD, a situation that could never lead to an arms race under full rationality and perfect information in an isolated game. Yet, if the missile fails in its role as a bargaining chip, substantial numbers may nonetheless be built.

Another important arms race-related behavior that can be placed

under the heading of strategic misrepresentation involves a pretense about capabilities rather than preferences. In the fall of 1957, the Soviet Union announced that it had orbited a satellite by means of an ICBM. In the months following this announcement, Khrushchev, through a combination of frequent public display of a limited number of missiles, clever rhetoric, and American ignorance, systematically misrepresented (and exaggerated) the actual ICBM capability of the Soviet Union, which was not close to that of the United States.[31] He subsequently attempted to exploit this "missile gap" to increase Soviet influence in various policy areas. Although he achieved only sporadic success, he did manage to alarm the American government about its supposed strategic weakness. The most significant evidence of this concern, of course, is to be found in the report of the Gaither Commission, which recommended large increases in defense expenditures as a response. President Eisenhower's attitude toward the balanced budget prevented him from fully embracing those recommendations, but he sharply accelerated existing ICBM programs (and the arms race)—an action that he would not have taken had the Soviets not misrepresented their capabilities.

B. IMPERFECT INTELLIGENCE

Imperfect intelligence can inspire an arms race that would not take place in the presence of perfect information, and can permit one to continue when it is "unjustified." Cobden and Hirst both argued that the British consistently overestimated the rate of French ship production and the size of the French fleet in each of the Anglo-French naval races, and that they responded with building programs of their own that wildly escalated whatever race may actually have existed.[32] In fact, on at least one occasion there is evidence that the size of the French fleet was actually declining while the British were convinced that they were in an arms race.[33] Similarly, a minor arms race was on the verge of taking place between France and Germany in 1875 because German intelligence reported heavy French purchases of cavalry horses. It took considerable effort on the part of the French ambassador to reassure the German government that its information was simply wrong.[34]

More recently, the adoption of "worst-case analysis"—in which rivals build weapons in response to the largest increases that their opponent could implement—has exacerbated the rate of arms escalation. In their

[31] Arnold Horelick and Myron Rush, *Strategic Power and Soviet Foreign Policy* (Chicago: University of Chicago Press, 1966).

[32] Cobden (fn. 5); F. W. Hirst, *The Six Panics* (London: Methuen, 1913).

[33] Cobden (fn. 5), 240.

[34] G. P. Gooch, *Franco-German Relations*, 1817-1914 (London: Longmans, Green, 1923).

discussion of how the Pentagon under McNamara assessed "how much is enough," Enthoven and Smith note the inclusion of generous assumptions about Soviet capabilities; although they were recognized as unlikely, "because such threats were conceivable and within Soviet technical capability, they were explicitly and systematically considered in the force planning process."[35] An even more outspoken statement of the tendency to assume that enemy capabilities are greater than they appear to be was offered by General Curtis LeMay in justifying the Air Force's strategic posture. "We have to have sufficient military power to knock out all of the targets we know he has, or all the weapons we know he has, and I would like to have a little cushion to take care of some that we might not know he has."[36]

Imperfect intelligence expands the range of games that can lead to arms races by raising the possibility that one side will think the other side has defected even though this may not have occurred. If we retain the requirement that DD > CD so that defection (or apparent defection) on the part of one side will cause the other side to defect as well, three additional games become potential sources of arms races.

10. CC > DD > DC > CD
11. CC > DC > DD > CD   (Stag Hunt)
12. CC > DD > CD > DC

These three games are not often analyzed; under full rationality and perfect information, each should inevitably lead to mutual cooperation because both parties will observe that this outcome is optimal for each of them individually and collectively. Moreover, all three games are cooperation-stable in the sense that there is no incentive for either party to defect from a cooperative solution.

The first and third bi-stable games (10 and 12) are not particularly useful because they require the unlikely circumstance that one party prefers mutual defection to defecting while the opponent cooperates. Stag Hunt, however, may frequently be an accurate model of actual arms races. It describes, for example, the classic situation in which a state acquires a weapon it would not have built on its own initiative because it believes that the rival is building it.[37]

[35] Alain Enthoven and Keith Smith, *How Much is Enough?* (New York: Harper & Row, 1971), 178.
[36] Quoted in William Kauffman, *The McNamara Strategy* (New York: Harper & Row, 1964), 284.
[37] Jervis (fn. 1).

C. PROBLEMS OF INTERPRETATION

The ideologies of decision makers and the experiences they have in operating in the international system inevitably color the way they process information as well as their vision of what arms race game they are engaged in. Whenever there is uncertainty about the likely impact of an action or the significance of another state's behavior—which is almost all the time—these factors play a major role in determining how they will assess the situation. Ideologies function like a prior distribution in conditioning assessments of probability under uncertainty. The difficulty, of course, is that there is no guarantee that this experience or ideologically guided vision is accurate. In terms of its effect on the games that can generate arms races, the impact of experience and ideology is identical with that of poor information; all create the possibility that an arms race will emerge from a Stag Hunt.

Instances where experience has colored national decision makers' estimates of the game in which they were engaged, and increased their tendencies to intensify the arms race, are plentiful. U.S. attitudes toward arms control immediately after World War II were heavily conditioned by the ultimate failure of the Washington Treaty to prevent Japanese arms increases. Secretary Forrestal opposed proposals for sharing the secret of making the atomic bomb with the Soviet Union in large part because he felt the Soviets were Asiatic in their mentality and hence, like the Japanese, could not be trusted to live up to arms limitation treaties.[38] Secretary Byrnes's position was not appreciably different: "America scrapped battleships, Japan scrapped blueprints. America will not again make that mistake."[39]

Experience is one source of problems of interpretation, but it is not necessarily the most important one. Ideology is a close rival. The available evidence suggests, for example, that John Foster Dulles's ideology was such that he had difficulty recognizing cooperative behavior on the part of the Soviet Union. Thus, in 1956, when the Soviets announced a reduction of 1.2 million in the size of their armed forces, Dulles, when questioned by reporters, emphasized the lack of trustworthiness of Soviet leaders. Similarly, Dulles greeted apparently conciliatory gestures by Stalin's immediate successors as the "tricks of rotten apples."[40] Whether Dulles's interpretation of these specific Soviet acts is correct or not is

---

[38] Ernest R. May, *"Lessons" of the Past: The Use and Misuse of History in American Foreign Policy* (New York: Oxford University Press, 1973), 33.

[39] *Ibid.*, 35.

[40] David Finlay, Ole Holsti, and Richard Fagen, *Enemies in Politics* (Chicago: Rand McNally, 1967), 60.

unimportant. The point is that Dulles's ideology made it difficult for the Soviets to communicate interest in initiating an end to any facet of the arms race if they so desired. In this context, an ideology that inspires the judgment that the rival has defected performs precisely the same function as imperfect intelligence that suggests that the rival is still "racing" when in fact it is not. We shall see that problems of interpretation can cause as much difficulty in sending a cooperative signal as they can in receiving one.

D. PROBLEMS OF CONTROL

Control problems contribute to the continuation of arms races whenever decision makers who intend to signal their willingness to cooperate (or even to defect in order to punish defection) cannot do so because other parts of the government are not responsive to their instructions or desires. In the midst of the Cuban missile crisis, President Kennedy became aware of the fact that his earlier orders to remove a number of obsolete and vulnerable (perhaps even provocative) Jupiter missiles in Turkey had not been followed. He again gave direct orders for their immediate removal, but such were the control problems that the initial result was nothing more than an examination of the options for defusing them.[41] What were the Soviets to make of all this? Were they dealing with a government willing to help them save face on the Cuban missile issue, or did the Americans intend to be provocative? For that matter, what were they to make of the Senate's failure many years later to ratify SALT II?

The control problem goes beyond the straightforward issue of whether or not direct orders are executed. As Halperin has pointed out, a well-timed leak to the press can significantly damage any policy that does not have wide support, and negotiations still in progress can frequently be destroyed by an untimely indiscretion.[42] Perhaps it is this realization that prompts national leaders to make use of personal envoys and summit meetings at critical stages in arms talks. Although the conventional explanation is that it is an adaptive response to overload in the regular channels of communication, it could also be due in some measure to a realization that the stakes are too high to risk a failure of control.

Uncertainty complicates the application of unilateral strategies in sev-

[41] Graham Allison, *The Essence of Decision* (Boston: Little, Brown, 1971), 142. An interesting line of inquiry challenges this account of the missiles in Turkey; see Barton Bernstein, "The Cuban Missile Crisis: Trading the Jupiters in Turkey?" *Political Science Quarterly* 95 (Spring 1980), 97-125.

[42] Morton Halperin, *Bureaucratic Politics and Foreign Policy* (Washington, DC: Brookings Institution, 1974).

eral ways. First, as we have already noted, if the rival is motivated more by ambitions to extend its power than by security concerns, then the unilateral steps the first state takes in order to feel more secure may only inspire still greater efforts on the part of the rival to exert its influence. The Maginot Line and the British fighter force both led the Germans to escalate their arms production and their tactical innovations. No doubt one of the reasons why the employment of unilateral strategies to promote cooperation is so rare is that governments are usually suspicious that their rival is motivated less by security than by the desire for greater policy leverage.

Second, the basic character of a particular weapon system or alliance is frequently a matter of interpretation. An example of this difference in interpretation can be found in the World Disarmament Conference held in Geneva in 1932. There it was proposed that offensive weapons be outlawed, while defensive weapons be allowed to remain in existence. Although they were able to agree that heavy tanks were clearly offensive weapons, the states that were represented at the conference disagreed about the character of *every* other weapon. The aircraft carrier, for example, was viewed to be offensive by the Japanese but defensive by the Americans.[43] Similar difficulties occurred in the different interpretations of frontier fortifications by the German and French governments in the 1930s. To the French, who by 1932 had invested a considerable amount in the Maginot Line, frontier fortifications were defensive. The German government, however, saw the region behind the Maginot Line as an area where troops could be concentrated for an attack on Germany. This view probably contributed to the construction of the fortresses on the Siegfried Line, which were placed directly opposite those of the Maginot Line.[44]

Although many alliances are intended to be defensive, they are not necessarily viewed as such by prospective opponents. The Anglo-French Entente of 1904 was created as a way of stabilizing Anglo-French relations in troublesome policy areas. Yet by 1905, the Germans saw the alliance as a design to threaten several of their interests, most notably their commercial interests in Morocco.[45]

The impact of the other complicating factors on the success of unilateral strategies is smaller than the impact of interpretation and often interacts with it. The quality of intelligence information can obviously

[43] George Quester, *Offense and Defense in the International System* (New York: Wiley, 1977).
[44] Anthony Kemp, *The Maginot Line: Myth and Reality* (London: Frederick Warne, 1981).
[45] Taylor (fn. 18), 404.

be important. The more open a government is about the character of its weaponry, the less likely it is that the worst-case scenario will dominate the other state's response. If a unilateral strategy is being pursued to mitigate an arms race, it is essential that the other side be convinced that any ABM system is useful only for silo defense, or that the work taking place at the missile silos is intended merely to harden them rather than to install multiple warheads.

Before exploring the impact of problems of information and control on tacit bargaining strategies, as represented by Tit-for-Tat, it is useful to distinguish the merits of reciprocity as an evolutionary strategy from its virtues as a strategy to end or reduce the intensity of a single arms race. As an evolutionary strategy, it is potentially applicable to the long-run development of everything from single-celled organisms to civilizations, and derives its strength from success or failure over many iterations. As a strategy to end a particular arms race, however, it depends on the expectations it creates about the probable consequences of arming or not arming to reduce the net benefit of defection or to transform the game from a Prisoners' Dilemma into something more benign. Not surprisingly, however, these expectations can be influenced by our set of complicating factors in such a way that the effectiveness of the Tit-for-Tat strategy is degraded. It should be noted that in Axelrod's experiments, Bayesian routines performed poorly;[46] nevertheless, expectations would still seem to be of great importance for tacit bargaining among states, for at least two reasons. First, a state faces another state with much more accurately established priors than if it faced a computer program of potentially arbitrary design. Second, a state has a much greater variety of information to condition on during the arms race than the computer program, which could only use the sequence of past defections and cooperations.

To appreciate the extent to which the effectiveness of Tit-for-Tat can be diminished by problems of information and control, it is helpful to construct a simple model. First, consider the problem of control. For reasons already discussed, when decision makers intend to perform a certain action X, such as taking missiles out of Turkey, there is only a certain probability (call it c) that X will be correctly implemented. Second, there is the problem of information flow. The action X that was initiated by the first party may or may not be accurately perceived by the second party. For example, the number of missiles put into or taken out of operation may be over- or under-estimated by the competing state.

[46] Axelrod (fn. 2), 192ff.

Third, the action X that may have been intended as cooperation (defection) may not be interpreted as such by the other state. This problem is especially acute when initiatives can potentially take place over a large number of weapon systems. If one side stops building ships but raises the rate of tank production, or slows its rate of constructing missile-carrying submarines but slightly increases cruise-missile production, will that be perceived as cooperation or defection? Can the initiating state be certain how such actions will be interpreted? Can the other state be certain how such actions were intended?

These problems may be hard to distinguish in practice. When information is poor, what one supposes the opponent to have done is heavily dependent on prior beliefs, conditioned by previous interactions. For this reason—and to keep the present discussion from becoming too complex—we will model problems of information and of interpretation together. If the opponent has done X, intended to be cooperation, let the probability that it will be perceived and interpreted as cooperation be q_1. If the opponent has done Y, intended to be defection, let the probability that it will be accurately perceived be q_2. In reality, control, information, and perception are not dichotomous; they occur in many shades along a continuum. Neither are defection and cooperation binary choices. Nonetheless, the simple model in which one pretends that all these factors are dichotomous is useful in that it eliminates many unnecessary complications.

It is interesting to see how complex the analysis may become in the presence of these factors. Suppose that state A decides to cooperate. If the action is correctly implemented (an event of probability c), cooperation is perceived with probability q_1. If it is incorrectly implemented (an event of probability $1 - c$), cooperation is perceived with probability $1 - q_2$. Thus, the probability of state B perceiving cooperation if A intended to cooperate is given by the upper right-hand value in Figure 2. Similar calculations lead to the other entries in Figure 2.

In the traditional, full-information, complete-control version of the model, the parameters c, q_1, and q_2 are all equal to one. If each party begins with cooperation, then with perfect information under Tit-for-Tat, the outcome is always CC. We may represent this as a vector of probabilities for the four states, CC, CD, DC, DD of (1.0, 0.0, 0.0, 0.0). Let us assume, however, that $q_1 = q_2 = 0.99$. With this only slightly more realistic assumption of a 1 percent chance of misperception (there is every indication that most arms races are plagued with a far higher figure), the long-run probabilities of the four outcomes are (0.25, 0.25, 0.25, 0.25). This means that the game stabilizes with at least one party

FIGURE 2
PROBABILITIES OF ACTUAL RESPONSES BEING PERCEIVED IN
VARIOUS WAYS ACCORDING TO ONE MODEL OF INFORMATION AND CONTROL

A's Intention	Chance that	
	B perceives Cooperation	B perceives Defection
Cooperate	$cq_1 + (1-c)(1-q_2)$	$c(1-q_1) + (1-c)q_2$
Defect	$c(1-q_2) + (1-c)q_1$	$cq_2 + (1-c)(1-q_1)$

defecting 75 percent of the time. Axelrod's results for a misperception variant of his tournament show it to be still superior to the other strategies entered, but those other strategies were designed under the assumption of perfect information.[47]

One reason for the sub-optimal outcome for Tit-for-Tat under conditions of imperfect information is that each side was too ready to defect after the apparent previous defection of the opponent. In this case, a little more patience would seem to be called for. A possible extension of this class of strategies is probabilistic Tit-for-Tat, in which the opponent's cooperation is followed by our cooperation with probability s and by our defection with probability $1 - s$, and in which the opponent's defection is followed by defection or cooperation with probabilities s and $1 - s$, respectively. When the opponent appears to have defected, we will defect with some probability less than 1, thus being somewhat more patient. To see how this works, let us assume that both countries have perfect control and that $q_1 = q_2 = 0.95$. Suppose that both countries decide to behave more forgivingly, treating a defection as a defection only part of the time: they behave as if q_2 were less than 0.95, say 0.5. In that case, the long-run probability vector is (0.83, 0.08, 0.08, 0.01), a much more satisfactory outcome. Even if only one country behaves forgivingly while the other pursues Tit-for-Tat, the outcome is (0.74, 0.10, 0.14, 0.02).

Axelrod has suggested that retaliation in Tit-for-Tat should be less than 1 in order to alleviate the consequences of a single defection, which under Tit-for-Tat with perfect information and control can lead to endless rounds of echoing or mutual defection.[48] It might be wisest to

[47] Ibid., 182ff.
[48] Ibid., 138 and 186ff.

combine this tactic with probabilistic Tit-for-Tat to produce a strategy that is, in effect, a probability distribution on the unit interval (where $0 =$ defection and $1 =$ cooperation) and values between 0 and 1 represent partial defection.

The same consequences of imperfect information hold true in an evolutionary sense. Suppose that conditions are such that $q_1 = q_2 = 0.95$. Under Tit-for-Tat (TFT), the state will cooperate on the first move and later do whatever the opponent is perceived to have done on the previous move. Modified Tit-for-Tat (MTFT) is similar except that it is more forgiving, behaving as if $q_2 = 0.5$. Using Axelrod's payoff matrix, the long-run payoff of TFT vs. TFT is 2.25, the payoff of TFT facing MTFT is 2.94, the payoff of MTFT vs. TFT is 2.74, and the payoff of MTFT vs. MTFT is 2.90. This means that a large population of TFT can be successfully invaded by MTFT so that TFT is not collectively stable under imperfect information. Of course, MTFT is not stable either; in equilibrium and in large populations, equal payoffs would be obtained if the population were 82.5 percent MTFT and 7.5 percent TFT. No doubt there is a strategy superior to either of these, but this remains a problem for future research.

The foregoing should not, of course, be taken as an indictment of tacit bargaining in general. Hirst credits a policy of reciprocity with having brought about the longest period of cooperation between the British and French during the 19th century.[49] In addition, Schelling has noted the emergence of several instances of arms restraint between the United States and the Soviet Union which were apparently achieved on the basis of tacit bargaining.[50] Nonetheless, it is important to recognize that tacit bargaining strategies can be very susceptible to problems of control, information, and interpretation. In a given context, the best solution may be to employ a Modified Tit-for-Tat, but an equally practical alternative may simply be to take elaborate measures to minimize control and information problems by bypassing normal channels, employing hotlines, and so forth. If these problems are effectively dealt with by either method, tacit bargaining strategies can prove invaluable in providing an opponent with incentives to cooperate in Prisoners' Dilemma or Stag Hunt.

The relationship between negotiation and problems of information and control is more indeterminate than was the case with the strategies

[49] Hirst (fn. 32), 39.
[50] Thomas Schelling, "Reciprocal Measures for Arms Stabilization," in Donald Brennan, ed., *Arms Control, Disarmaments, and National Security* (New York: George Braziller, 1961), 175. More recently, Russell Leng has shown the effectiveness of reciprocity-based tactics in Soviet-American relations. "Reagan and the Russians: Crisis Bargaining Beliefs and the Historical Record," *American Political Science Review* 78 (June 1984), 338-55.

of unilateral action and tacit bargaining. In part this is because one of the great strengths of negotiation lies in its ability to cope with the same information problems that can undermine the other strategies. The fact that the United States and the Soviet Union place a different value on throw-weight versus number of missiles, or that the Germans and the British placed a different value on numbers of battleships versus broadside weight can play havoc with a strategy of tacit bargaining, since neither side can fully understand what the other would view as reciprocity. Yet the difference in utilities that can be the source of so many problems of interpretation is precisely what negotiation can exploit in order to create an agreement.

The strategic misrepresentation that attends negotiation can occasionally create problems of information and interpretation that lead to still greater arms escalation. Before the onset of British mediation in the naval race of 1894-1902 between Argentina and Chile, the exaggerated claims and aggressiveness employed by both countries in their intermittent negotiations seemed to inspire ever greater hostility in the other party, and consequently increased arms competition.[51] Moreover, as we noted earlier, the use of bargaining-chip strategies, such as threatening to put in a weapons system that is really undesired in order to gain concessions from the other side, can lead to intensified arms competition that would never have taken place without negotiation.

The strategy of negotiation can also intensify an arms race by providing both sides with an incentive to bargain from as strong a position as possible. It is not difficult to deduce the arms-race implications of Richard Nixon's philosophy that no state should ever let its president go to the bargaining table as the leader of the second-strongest state in the world. Such states will prepare for peace talks by escalating the arms race, and probably feel compelled to maintain this pace if the talks collapse. Even more perversely, each will assume that the other will adopt this strategy—and this will increase the rate of arms acquisition still further. Note that this is different from the typical bargaining-chip strategy in that both states are arming rather than threatening to arm, and it is a problem that plagues neither unilateral action nor tacit bargaining.

IV. CONCLUSION

Both the "game" that underlies an arms race and the conditions under which it is conducted can dramatically affect the success of any strategy

[51] Robert Burr, *By Reason or Force: Chile and the Balancing of Power in South America, 1830-1905* (Berkeley: University of California Press, 1965).

designed to end it. If antagonists have Deadlock or Deadlock-type pref-
erences, only negotiation can provide the means by which cooperation
might be strategically achieved; even this possibility depends on the pre-
existence of linkages between the arms race and the rules that govern
interactions among the antagonists on other issues (for example, trade).
Behaviors that resemble unilateral strategies may be employed and may
reduce the intensity of the arms race, but they are not motivated by a
desire for cooperation; they will only be adopted if they are optimal
from the standpoint of security. Tacit bargaining strategies will not be
employed at all. The possibility that Deadlock-driven arms races may
be more common than is generally assumed helps to explain both why
the use of cooperative strategies is relatively rare even in the set of arms
races that have ended cooperatively, and why the cooperation that has
occurred has often been prompted by an economic change and third-
party intervention.

If an arms race can be described as a Prisoners' Dilemma (at least
from the standpoint of preference orderings) conducted under conditions
of perfect information, the outlook for cooperation and for the success
of all three cooperative strategies is more sanguine—provided certain
critical assumptions turn out to be correct. Assuming that both sides are
driven by security needs rather than by a desire for policy influence—
an assumption that may well be false—unilateral strategies may promote
cooperation by reducing the benefit of an aggressive arms policy on the
part of the rival, and the cost of subsequent cooperation for the state
employing it. Assuming the benefits of future cooperation (the shadow
of the future) to be great enough when compared to the potential costs
of falling behind in the arms race—an assumption that states often seem
reluctant to make—the tacit bargaining strategy of Tit-for-Tat can be
an effective way of ending an arms race. Negotiation can succeed in a
Prisoners' Dilemma environment by forging a link between the arms
race and other games or issues. As long as the benefits of this linkage
can be withdrawn to punish defections, there will be an incentive for
continued cooperation.

When we drop the assumption of perfect information and admit the
existence of problems of information and control that are often present
in actual—as opposed to hypothetical—arms races, each of the strategies
faces additional limitations. In the case of unilateral strategies, the biggest
threat to success comes from problems of interpretation. What is viewed
as defensive by a government interested in initiating the end of the
security dilemma may not be so viewed by its rival. In the case of tacit
bargaining, the problem of interpretation also looms largest, but those

of information and control can be important too. Modest reductions in the probability that an action believed by the initiating state to be co-operative will be correctly executed, observed, and interpreted can lead to startlingly poor results for a tacit bargaining strategy that might succeed admirably under optimal conditions. In the case of negotiation, strategic misrepresentation can so muddle the character of interaction that what could have been a golden opportunity for cooperation is lost.

An understanding of the sources of uncertainty that frustrate the application of cooperative strategies also helps us to understand some of their special strengths. If the problems that are internal to their application can be overcome, they can help uncover critical facts about the nature of the arms race that is taking place, and defuse that com-ponent of it which derives from misperception rather than from fun-damental policy disagreement. Unilateral strategies can provide a par-ticularly safe avenue for trying to determine the motivations of an adversary and for identifying the game being played. If a defensive alliance or weapon system leads to far greater weapons activity than could reasonably stem from misperception alone, the game that the two states are playing is not Stag Hunt, and the rival state is motivated by policy influence goals as well as by security considerations. Simple tacit bargaining strategies conducted in a highly iterative setting with quick responses to small instances of cooperation and defection can provide a particularly good environment in which states can learn about each other's preferences and willingness to cooperate. Negotiation that is not itself plagued with misrepresentation can occasionally reveal information that, even in the absence of an agreement, leads both sides to believe that the intensity of an arms race can be safely reduced.

Finally, it might appear that the best way for a state interested in promoting cooperation to cope with the difficulties associated with each individual strategy is to pursue all of them simultaneously. Although certainly desirable from the standpoint of exploiting the ways in which they complement each other, this joint strategy suffers from the fact that they are often mutually exclusive. Negotiators may wish to maxi-mize their bargaining position and the credibility of their threats by increasing their weapons production and development prior to negoti-ations, but this precludes the give-and-take of a Tit-for-Tat bargaining strategy. When the weapons are offensive in character—as they often are in order to provide the rival with the maximum incentive to negotiate a halt to the race—unilateral strategies are difficult to implement. Tit-for-Tat is also somewhat inconsistent with unilateral strategies, since the "punishment" of an increase in offensive weapons on the part of

the rival may necessitate a similar action in response. It is true, of course, that a unilateral strategy can be continued even as this punishment is being carried out, but the potential for misinterpretation will be dramatically increased by the heightened climate of tension that the act of punishment may produce. The existence of these incompatibilities makes it all the more important to understand their relative effectiveness under different circumstances.