

## Trust But Verify: The Production of Information in Arms Control Treaties and Other International Agreements

Kenneth W. Abbott

Follow this and additional works at: <http://scholarship.law.cornell.edu/cilj>



Part of the [Law Commons](#)

---

### Recommended Citation

Abbott, Kenneth W. (1993) "Trust But Verify: The Production of Information in Arms Control Treaties and Other International Agreements," *Cornell International Law Journal*: Vol. 26: Iss. 1, Article 1.  
Available at: <http://scholarship.law.cornell.edu/cilj/vol26/iss1/1>

This Article is brought to you for free and open access by Scholarship@Cornell Law: A Digital Repository. It has been accepted for inclusion in Cornell International Law Journal by an authorized administrator of Scholarship@Cornell Law: A Digital Repository. For more information, please contact [jmp8@cornell.edu](mailto:jmp8@cornell.edu).

## “Trust But Verify”:\*\* The Production of Information in Arms Control Treaties and Other International Agreements

### Introduction

This article utilizes the lens of rationalist international relations (IR) theory to examine an important class of provisions in international agreements, those governing the production of information. Rationalist IR theory<sup>1</sup> assumes that states<sup>2</sup> act as rational entities pursuing their national interests as they see them.<sup>3</sup> In situations of interdependence, the theory suggests, states will, and should, tend to design their international agreements and institutions to address the particular strategic situations<sup>4</sup> in which they find themselves.<sup>5</sup>

---

\* Professor of Law, Northwestern University School of Law; B.A., Cornell University; J.D., Harvard Law School.

I would like to thank Ron Allen, Charles Lipson, Duncan Snidal, Alex Wendt, and the student and faculty participants in seminars at the Harvard Department of Government, Northwestern Law School, and Yale Law School for helpful comments and advice.

\*\* President Ronald Reagan, when asked whether he trusted the peaceful overtures being made by General Secretary Gorbachev, reported that in arms talks with Gorbachev he had used an old Russian saying “*Doveryai, no proveryai*,” translated as “trust but verify.” N.Y. TIMES, June 12, 1987, at A12. The phrase became a classic encapsulation of American arms control policy.

1. See Kenneth W. Abbott, *Modern International Relations Theory: A Prospectus for International Lawyers*, 14 YALE J. INT’L L. 335 (1989).

2. Rationalist IR theory commonly focuses on states as the most important actors in international politics. See *id.* at 346. This focus seems appropriate in the present context, since states are by far the predominant participants in international agreements.

3. In addition to the assumptions of rationality and egoism, rationalist theory often assumes that states act as unitary entities. See *id.* at 348-51.

4. A “strategic situation” is one involving a relationship of interdependence among a relatively small number of actors. When states—like chess players, oligopolists, or opposing armed forces—are involved in a strategic interaction, each state’s actions affect the fortunes of others as well its own, and the best course of action depends on what others may do. In international politics, a number of background

This insight, which might be called the "rational design hypothesis," has important implications for IR and for the study of international law. For one thing, it suggests that scholars can reason backward from the provisions of international agreements and the procedures and institutions they establish to conclusions about the strategic relationships of the parties to those arrangements. In addition to its importance for understanding international politics, the rational design hypothesis shows how international lawyers, expert in the interpretation of international agreements, can contribute to the IR research agenda.

Conversely, the rational design hypothesis suggests that scholars can reason forward from a theoretical understanding of particular issue areas to richer explanations of the meaning and function of international agreements, procedures and institutions.<sup>6</sup> By the same token, such structural analyses may make important normative contributions by suggesting provisions and mechanisms likely to contribute to strong and stable agreements in particular strategic situations.<sup>7</sup>

Many rationalist analyses have ignored the role of information in strategic interactions among states. Information problems have been assumed away by the use of models in which the participants have complete information about the structure of their relationship and the actions of others within it, or in which certain classes of information are simply irrelevant.<sup>8</sup> These approaches, however, ignore a central reality of international politics, its "pervasive uncertainty."<sup>9</sup> Recent IR scholarship, in contrast, has devoted increasing attention to problems of information.<sup>10</sup>

---

factors, such as the absence of central authority, also help define particular strategic situations.

5. Duncan Snidal, *Coordination versus Prisoners' Dilemma: Implications for International Cooperation and Regimes*, 79 AM. POL. SCI. REV. 923 (1985). See also Arthur A. Stein, *Coordination and Collaboration: Regimes in an Anarchic World*, 36 INT'L ORG. 299, 311-16 (1982).

6. For a discussion of the value of rich description, as compared with efforts at scientific explanation and prediction, see Abbott, *supra* note 1, at 353-54.

7. For those more scientifically inclined, the study of the structure of international interactions may produce testable predictions of the types of provisions likely to appear in agreements among states. See Duncan Snidal, *The Game Theory of International Politics*, 38 WORLD POL. 25, 27 (1985). Duncan Snidal has demonstrated the considerable differences one would expect to find in agreements designed to deal with situations resembling the Prisoners' Dilemma game and those in which states seek only to coordinate their actions. See Snidal, *supra* note 5.

8. See Jonathan Bendor & Thomas K. Hammond, *Rethinking Allison's Models*, 86 AM. POL. SCI. REV. 301, 303, 306 (1992); GEORGE W. DOWNS & DAVID M. ROCKE, *TACIT BARGAINING, ARMS RACES, AND ARMS CONTROL* 80 (1990); PETER C. ORDESHOOK, *MODELS OF STRATEGIC CHOICE IN POLITICS* 1-3 (Peter C. Ordeshook ed., 1989); Jonathan Bendor et al., *When in Doubt . . . Cooperation in a Noisy Prisoner's Dilemma*, 35 J. CONFLICT RESOL. 691, 692 (1992). For further discussion of assumptions that make information irrelevant, see *infra* text accompanying notes 31, 41, 48, 65, 84-85.

9. See Bendor & Hammond, *supra* note 8, at 303.

10. *Id.* at 306; ORDESHOOK, *supra* note 8, at 1.

This article joins the rational design hypothesis with the new emphasis on information by examining the kinds of informational arrangements that states make when they enter into agreements in conditions of interdependence. The article achieves greater concreteness than most other recent work by analyzing the provisions governing the production of information in a set of actual international agreements: the major arms control treaties of the past 30 years.<sup>11</sup> These agreements form a particularly appropriate subject for this kind of analysis, as problems of information have been at the heart of arms control policy since the beginning.

This article uses models drawn from elementary game theory, to depict the strategic settings of state interaction. Game models, designed precisely to highlight the strategic structure of interdependent relationships, represent ideal tools for studying the rational design of international agreements.<sup>12</sup> The use of game models sometimes meets resistance because it entails the sacrifice of much detail and nuance. In order to pursue the sort of analysis attempted here, however, one must rely on relatively spare and abstract models. A fully descriptive account of virtually any interaction would obscure its strategic structure in a profusion of facts.<sup>13</sup>

This article uses the well-known Prisoner's Dilemma (PD) game as its primary model of state interaction, but it also utilizes the game known as Stag Hunt (SH) or Assurance. Both are widely seen as representing many situations in international politics.<sup>14</sup> PD and SH are mixed-motive games: each state in such an interaction has incentives to cooperate with the others, but each may also have incentives to act independently, affecting the others negatively.<sup>15</sup> States may enter into an agreement to manifest and reinforce their cooperation,<sup>16</sup> but the same conflicting incentives will remain at work. Each state will have reason to

---

11. For a listing of the treaties considered in the article, see *infra* text accompanying notes 95-103.

12. For discussions of the value of game theory in analyzing international and other social interactions, see Abbott, *supra* note 1, at 357-58; RUSSEL HARDIN, COLLECTIVE ACTION 22-23 (1982).

13. According to Bendor and Hammond, game models are used "to conserve the intellectual resources of scholars and to focus these resources on analyzing the strategic relations among nations." Bendor & Hammond, *supra* note 8, at 307.

14. For further discussion of these games, see Abbott, *supra* note 1, at 358-60, 368-71.

15. There are 78 different 2 X 2 games, the simplest and most common form, in which two players must each choose between two moves. Out of this total, there are six games in which (a) each player prefers mutual cooperation to mutual defection, providing an incentive for agreement, and (b) each player prefers that if one player is to defect, it be the one to do so, providing a basis for conflict. These six games, including PD and SH, as well as Chicken and complex coordination games, have been the subject of most research in the application of game theory to social, political and economic cooperation.

16. An agreement will increase the certainty of iteration, raise the costs of breach or withdrawal, and spell out the parties' obligations; it may also provide procedures for interpreting ambiguities, authorize retaliatory measures, and the like. See Abbott, *supra* note 1, at 365-68; *infra* note 40.

observe the agreement, but each may also have reason to breach it, openly or by cheating or shirking, to withdraw from it, or otherwise to "defect." These incentives are of two general kinds: offensive, based on a desire to obtain a superior result for oneself, and defensive, based on a desire to avoid being hurt by the defection of others.

In mixed-motive situations like these, information regarding the structure of the interaction, the incentives perceived by other states, and the compliance of others with their obligations will be crucial to international cooperation. Information on compliance is particularly important, both for itself and for the light it can shed on other issues. States will be reluctant to enter into agreements without clearly defined mechanisms for the ongoing production of reasonably timely and reliable information on these matters. Such mechanisms (contained either within the agreements themselves or parallel to them) may determine the success of an agreement in practice.

This article identifies two basic strategies by which information is produced: verification and assurance. Verification—a term familiar from discussions of arms control, such as the quotation from former President Reagan in the title—involves each party's use of its own efforts and resources to seek out the information it requires from others.<sup>17</sup> States relying on pure verification assume that neither the other parties to an agreement nor any central authority will assist in the gathering of information; indeed, they may presume that others will try to frustrate their efforts by concealing relevant facts or attempting to mislead. Verification techniques range from research in publicly available documents through external observation to espionage. States can, of course, utilize many of these techniques unilaterally, without any explicit arrangements. Often, however, agreed procedures—like the on-site inspections authorized in recent arms control agreements—can greatly increase the effectiveness of verification.

The strategy of assurance, a less familiar term, involves the use by each party of its own efforts and resources to gather and provide information about itself and its activities to others. By providing information, each party demonstrates to the others its continued compliance with the terms of an agreement, thus forestalling defensive defection. Like verification, pure assurance assumes minimal participation by the other parties.<sup>18</sup> Assurance techniques can range from simple certifications of relevant facts to the provision of otherwise confidential internal documents and physical evidence. As with verification, states can provide many forms of assurance unilaterally, as the need arises. Again, however, including specific arrangements in an agreement can often

---

17. States may also rely on international organizations or third states with greater resources to gather information about a target state. See *infra* text accompanying notes 116-27.

18. As with verification, however, international organizations or third states can play a role in the assurance strategy by receiving, assessing and distributing information.

increase the effectiveness of assurance, while supporting continued cooperation in other ways.

Each of these strategies may be utilized independently. They are, however, closely related, and states facing the complex incentives of PD and SH situations will often have reason to utilize both, or to combine them in various ways. Some of the most effective assurance devices in arms control agreements, for example, are carefully designed to mesh with external monitoring procedures; similarly, extensive verification regimes like on-site inspection are only workable because of the cooperation of the parties under observation. Such hybrids are typically the most effective mechanisms of all. In areas like arms control where hybrid mechanisms are common, verification and assurance can best be viewed as ideal types, useful heuristically for the analysis of complex informational arrangements.

Section 1 of the article analyzes the PD game, the various conditions under which it may be played, and the informational requirements of states involved in such interactions. Section 2 derives the strategies of verification and assurance from the incentives of the PD game and introduces the SH game to expand the analysis. Section 2 also explores the types of mechanisms that states might use to implement these strategies, focusing on those that require explicit *ex ante* recognition in international agreements. Section 3 examines in detail how the strategies of verification and assurance have been utilized in arms control agreements. The final section is a brief conclusion that suggests how verification and assurance techniques might apply to future arms control agreements.

## I. The Prisoners' Dilemma Game (PD)

### A. The Standard PD

PD is a positive-sum game, in which cooperation among the players can increase each player's welfare (its "payoff") as compared to independent action.<sup>19</sup> As noted above, however, PD is also a mixed-motive game, in which each party may find it advantageous to defect from cooperative arrangements, reducing the payoffs of others in the process.<sup>20</sup> The constant tension between the centripetal incentives to cooperate and the

---

19. It is widely understood that agreements are devices to capture the gains available in such positive-sum situations. For an example in the context of ordinary contracts among individuals and firms, see IAN R. MACNEIL, *CASES AND MATERIALS ON CONTRACTS* 1-5 (1971).

In situations resembling the game of Harmony, in contrast, no agreement is necessary, as each player's pursuit of its own interest automatically furthers the interests of others. For discussions of Harmony, see Abbott, *supra* note 1, at 354-57; ROBERT O. KEOHANE, *AFTER HEGEMONY: COOPERATION AND DISCORD IN THE WORLD POLITICAL ECONOMY* 51-52 (1984) (distinguishing harmony situations from true cooperation).

20. In contrast to mixed-motive games, games of pure coordination, see Abbott, *supra* note 1, at 371-72, are wholly harmonious, while constant-sum games like poker are wholly conflictual. Stag Hunt, under assumptions of complete information, can become a game of pure coordination. See *infra* text accompanying notes 84-85.

centrifugal incentives to defect creates an urgent need for information.<sup>21</sup>

PD is the most widely studied of all games. It is often taken to encapsulate, among other things, the reasons for establishing states with sovereign authority, the problem of providing collective goods, the difficulties of collective action in a diverse array of social settings, and the incentive structure of exchange relationships.<sup>22</sup> PD is also regarded as the model with the greatest power to elucidate the problems of conflict and cooperation in the decentralized international system.<sup>23</sup>

The game derives its name from the story of two prisoners. The police believe the two were accomplices in an armed robbery, but do not have enough evidence to convict them without a confession. The prisoners are separated, so they cannot communicate, and the prosecutor offers each of them the following deal: (a) if neither of you talks, we can only convict you of a minor offense, but each of you will still receive a year in jail; (b) if either one of you confesses fully, we will let that one off entirely, while the other will get ten years for armed robbery; or (c) if both confess, both will be convicted of armed robbery, but we will ask for a lenient sentence of six years.

The "normal form" representation of this story is the symmetrical 2 X 2 matrix shown in Figure 1.<sup>24</sup>

The cells in the matrix represent the factual outcomes resulting from the four possible combinations of player choices or "moves": in the prisoners' story, for example, the upper left represents mutual silence, the lower right mutual confession. The numbers in the cells represent the values these outcomes have for the players: their payoffs. The payoffs to Player I are given first. In Figure 1 the payoffs are cardinal numbers, corresponding to specific jail terms postulated by the prosecutor; here, unlike most PD models, the higher number represents a less desirable outcome. Cardinal payoffs greatly increase the power of game models, by revealing the intensity with which certain outcomes are

21. Some relevant information will be available in the ordinary course. Cf. R. Harrison Wagner, *The Theory of Games and the Problem of International Cooperation*, 77 AM. POL. SCI. REV. 330, 331 (1983) (unlike usual game theory assumption, states often have information about others' actions). The available information, however, will often be insufficient.

This article focuses on explicit arrangements for the production of additional information. As discussed further below, many versions of the PD game assume away any informational requirements.

22. See, e.g., HARDIN, *supra* note 12, at 25; MICHAEL TAYLOR, *THE POSSIBILITY OF COOPERATION* (1987); ROBERT AXELROD, *THE EVOLUTION OF COOPERATION* 3 (1984); Russel Hardin, *Exchange Theory on Strategic Bases*, 21 SOC. SCI. INFO. 251 (1982).

23. See, e.g., Hardin, *supra* note 22; TAYLOR, *supra* note 22; COOPERATION UNDER ANARCHY (Kenneth A. Oye ed., 1986); Abbott, *supra* note 1, at 360-62.

24. The 2 X 2 matrix, or normal form, in which two players each face two choices, is the simplest and most common representation of games of interdependence. The lessons of 2 X 2 games can often be extended to games with larger numbers of players, but the analysis may become considerably more complex. That extension is not explicitly made in this article, the main goal of which is to analyze the strategies of verification and assurance.

**Figure 1**

		Player II	
		Silence	Confess
Player I	Silence	1 year, 1 year	10 years, 0
	Confess	0, 10 years	6 years, 6 years

preferred to others, and by allowing for comparison of preferences over time and between games. In most cases, however, cardinal payoffs are impossible to identify; the most that can be specified is an ordinal ranking, perhaps with a rough estimate of relative intensity.<sup>25</sup> Figure 2 shows a more general PD matrix, with ordinal payoffs—from 4, the highest, to 1, the lowest—and a generic choice between “cooperation” (C) and “defection” (D).<sup>26</sup>

**Figure 2**

		Player II	
		C	D
Player I	C	3, 3	1, 4
	D	4, 1	2, 2

C = Cooperation                      D = Defection

In the standard “one-shot” version of the PD game—which assumes, among other things, that each player can make only one move and that the game is not repeated—each player is better off defecting,

25. For useful discussions of preferences, see Snidal, *supra* note 7, at 46-48; RUSSEL HARDIN, *MORALITY WITHIN THE LIMITS OF REASON* 53-55 (1988) (discussing implications for ethical analysis). The matrix in Figure 3 utilizes interval-level payoffs that estimate relative intensity.

26. It is worth observing that it is the prisoners' cooperation that is at issue in the PD game. Cooperation among persons accused of crimes, like that among oligopolists, is not socially desirable. The whole point of the prosecutor's strategem is to make cooperation difficult.

whatever the other player does. In game theory jargon, defection is a "dominant strategy." The dilemma is that, if each of the prisoners rationally follows its dominant strategy, both will confess, giving each the third-best outcome (a six-year sentence). If both could only keep silent, counter to their individual incentives, each would receive the second-best outcome (a one-year sentence). International analogues include arms races and competitive economic policies like strategic subsidies and optimal tariffs.<sup>27</sup>

Looking more closely at the PD incentives, one sees that each player receives the highest possible payoff by defecting while the other continues to cooperate; each player receives the lowest possible payoff if it continues to cooperate while the other defects. In fact, then, each player faces two separate incentives to defect. One is offensive: by defecting unilaterally, a player can obtain the best outcome, leaving the other (incidentally or by design) with the worst.<sup>28</sup> The second is defensive: by defecting, a player can avoid the worst outcome, the "sucker's payoff" for unilateral cooperation. In the PD game, these two incentives are inextricably linked. They are, however, logically distinct, and as such figure importantly in the analysis of mechanisms for the production of information.

If the players in a PD could communicate, as states normally can, they might well agree to cooperate once their options became clear. Their agreement would likely break down in practice, however, without some external agent (a government or a Mafia hit team) to enforce it; in international relations, of course, such agents are virtually nonexistent. Even if the players trusted each other fully,<sup>29</sup> offensive defection might sometimes be a temptation, and a defensive incentive to defect would thus remain.

Because cooperation is so difficult in the standard or one-shot PD, that version of the game is not centrally relevant to the study of international agreements. In any event, the one-shot PD is not a broadly accurate model of international relations. More realistic versions of the game will prove more fruitful.

---

27. See Abbott, *supra* note 1, at 360-62.

28. As this sentence suggests, the offensive incentive can be interpreted in two ways. If the players are concerned with absolute gains, they will tend to defect in order to obtain a high payoff. If they are concerned with relative results, or with minimizing the results of others, they will tend to defect because of the negative externalities that defection creates. A predatory strategy may contain elements of both approaches.

29. Trust might result from such factors as common ethnic or religious ties, inculcated ideals of solidarity, or the experience of prior cooperation. Similar factors can operate in international settings, but they are uncommon. For example, relations between the United States and Canada may benefit from a long history of cooperative interaction.

## B. Second-chance PD Games

In the standard PD game, neither player can change its move after learning what the other has done. As R. Harrison Wagner puts it, "[T]he crucial feature in the . . . story of the two prisoners is not that they were separated, but that once one had confessed while the other kept silent for a time, the latter was deprived of another opportunity to confess."<sup>30</sup> When only one move is possible, the twin incentives to defect are overwhelming. Since neither incentive depends on the other player's actual move, moreover, information about the conduct of others is irrelevant.<sup>31</sup>

If players have the ability to revise their moves in response to the actions of others, however, the situation is much different. First, as Wagner suggests, each player may be able to avoid the sucker's payoff; this should make them more willing to cooperate at the outset. More broadly, each can reward or penalize the other for its prior move, using strategies of conditional cooperation. For example, each player may reciprocate the other's moves tit for tat, continuing to cooperate as long as the other does, but responding to a defection with another defection.<sup>32</sup> So long as each state can recognize and respond to the other's actions with some degree of certainty, these possibilities tend to neutralize both incentives to defect.<sup>33</sup>

Figure 3—an "extensive form" representation of a PD game in which players move sequentially and can retract and revise their previous moves—demonstrates the possibilities for cooperation.<sup>34</sup>

Here Player I moves first, choosing between C and D; Player II makes its choice at the appropriate point on the decision tree; and so on. Each can change its move, beginning at the third decision point, until neither can improve its payoff further, that is, until an equilibrium is reached.<sup>35</sup> The most important branch on the tree shows Player I choosing C initially, Player II attempting to take advantage by choosing D, and Player I responding by changing its move to D. At this point, Player II might well change to C (aware that it can change back again), in the hope of initiating mutual cooperation. Just as important, if Player II could have anticipated Player I's response at the outset, it might well have chosen C in the first instance, making mutual cooperation an equi-

30. Wagner, *supra* note 21, at 331.

31. Information about the other players' incentives and other structural features of the game will be relevant, but full knowledge of these features is assumed in the standard PD. See *infra* text accompanying note 48.

32. Robert Axelrod and other theorists have identified this strategy as the key to cooperation in decentralized settings like international relations. See AXELROD, *supra* note 22, at 13; Abbott, *supra* note 1, at 365-66.

33. See Wagner, *supra* note 21, at 332-33.

34. Cf. PETER C. ORDESHOOK, *GAME THEORY AND POLITICAL THEORY: AN INTRODUCTION* 221 (1986) (applying an iterated PD model to the arms race).

35. In the standard PD, the only equilibrium is DD. That outcome is also an equilibrium, albeit an inferior one, in Figure 3. *Id.* at 221.

librium outcome.<sup>36</sup>

In principle, states are almost always able to retract or modify their earlier actions.<sup>37</sup> In practice, such reversals of field may be costly, and may encounter resistance within a state's bureaucracy, legislature or electorate.<sup>38</sup> States will only enter into a risky international agreement if they determine that they can recognize and respond to the actions of the other parties with reasonable certainty.<sup>39</sup> The ability to respond is largely an internal matter, though international agreements and institutions can make effective reactions more likely.<sup>40</sup> The ability to recognize also depends initially on national capabilities, but can be improved significantly by agreed mechanisms for the production of timely and accurate information.

### C. Iterated PD Games

In the standard PD game, there is only one interaction: it is assumed that the players will never again meet in the same situation. Like the inability to change moves, this assumption reinforces the dominance of defection and makes information about other players' actions largely irrelevant. If the interaction is repeated or "iterated," however, the situation again changes dramatically. So long as they can recognize each

---

36. Figure 3 can also be interpreted in terms of initial strategies. If each player follows the standard PD strategy of defecting whatever the other does, the result is an equilibrium at DD. If each follows the strategy of conditional cooperation, however, the result is an equilibrium at CC. Since the payoffs in the latter case are superior, conditional cooperation should be preferred by rational players. If one came to believe that one's next move would be the last, or next to last, however, defection would again become the dominant strategy. *Id.* at 221-22.

37. For example, states can institute subsidies or similar measures in response to predatory economic actions by others, as well as institute or speed up weapons programs or deployments in response to security threats by others. Both the initial policies and the countermeasures can then be reversed as the game is played out.

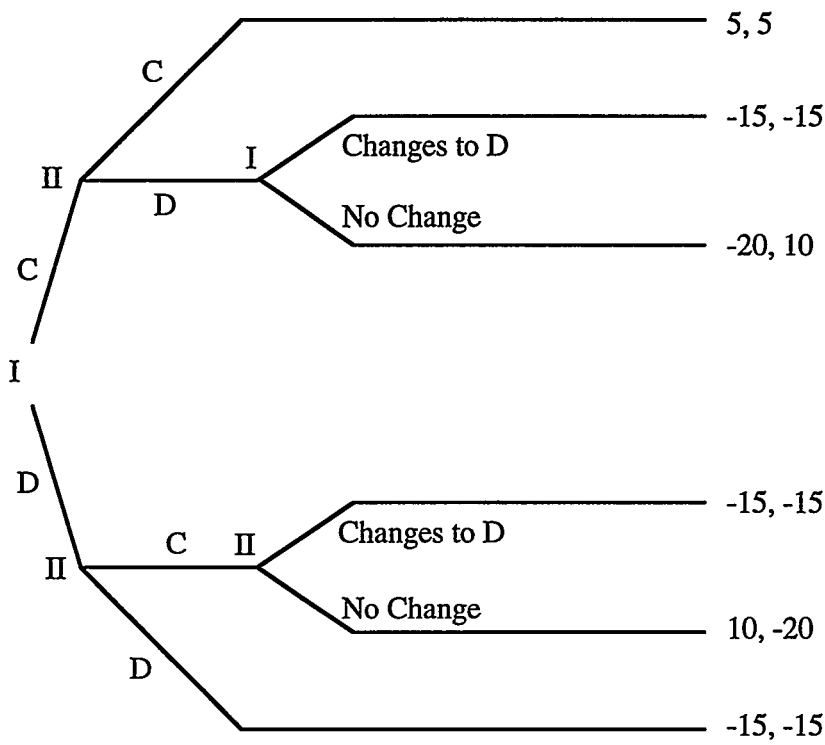
38. Rationalist IR theory generally proceeds on the assumption that states act as if they were unitary actors. *See supra* note 3. This assumption is normally necessary for strategic analysis, and is quite reasonable when dealing with arms control and similar issues. *See* DOWNS & ROCKE, *supra* note 8, at 92-97 (endorsing unitary actor assumption but favoring a more complex model than 2X2 PD). The assumption can, however, be relaxed when the internal determinants of state behavior become relevant. *See* Abbott, *supra* note 1, at 349.

A "strong state," in which the government has considerable autonomy from interests in the larger society, will be able to respond more freely than a "weak state." For the leading discussion of strong and weak states, see STEPHEN D. KRASNER, *DEFENDING THE NATIONAL INTEREST: RAW MATERIALS INVESTMENTS AND U.S. FOREIGN POLICY* (1978).

39. Perhaps surprisingly, a relatively low probability of recognition and a tit-for-tat response can be sufficient to induce cooperation, unless the incentives for defection are extremely strong or those for cooperation extremely weak. *See* ORDESHOOK, *supra* note 34, at 221 (illustrating an extensive form game with interval-level preferences in which 20% probability of tit for tat response will induce cooperation).

40. For example, agreements may authorize withdrawal or the use of particular countermeasures in specified circumstances. International institutions are useful in coordinating reactions, thus avoiding free-rider problems, when more than one state is involved. *See* Abbott, *supra* note 1, at 366-68.

Figure 3



I = Player I, II = Player II, C = Cooperation, D = Defection

other's prior moves, players can engage in conditional cooperation, rewarding or punishing those moves in subsequent iterations.

Ex ante, this possibility weakens the offensive incentive to defect: even if a party could reap a high payoff in the short run by defecting unilaterally, it might be forced to endure a lengthy period of non-cooperation, with inferior payoffs.<sup>41</sup> Any weakening of the offensive incentive to defect, of course, weakens the defensive incentive as well. Iteration also strengthens the incentive to cooperate: a cooperative move now may initiate years of beneficial mutual cooperation.<sup>42</sup> Relatively certain iteration tends to transform PD into a more cooperative game.<sup>43</sup>

In many areas of international relations, states can reasonably anticipate that their interactions will continue indefinitely. The major

41. The influence of such future events on current actions will depend on a variety of factors. See *infra* text accompanying notes 60-61.  
42. As with the revision of moves, however, if a particular interaction is believed to be the last, or close to the last, the dominance of defection is likely to take over.  
43. See *infra* text accompanying note 89.

nations of Europe, for example, can expect to deal with each other on issues like economics, security, and human rights well into the future; the same is true for the United States vis-à-vis the European Community, Canada, and Mexico. Although conditions have now changed drastically, the United States and the Soviet Union reasonably anticipated an indefinite relationship throughout the arms control era.<sup>44</sup> In circumstances like these, there will be many opportunities for tit for tat responses.

In addition to iteration over time, different issue areas are often linked, formally or informally, in the minds of government officials and their publics. When this is so, PD relationships are in essence iterated horizontally, across issue areas. Horizontal iteration can be especially important when agreements have a defined term. Last-minute cheating on an arms control agreement, for example, can be expected to sour chances for future cooperation on economic assistance or trade.<sup>45</sup> More generally, defections may contribute to an overall reputation for untrustworthiness, making it more difficult for the defecting state to enter into desired agreements, even with entirely different states.<sup>46</sup>

Strong expectations of iteration will be an important consideration in decisions by states to enter into international agreements. As in second-chance games, states must also have reasonable confidence in their own ability, and that of others, to recognize moves and respond to them appropriately. Here again, states will frequently seek to establish agreed mechanisms for the production of timely and reliable information.

#### D. The Demand for Information

In order to analyze more completely the demand for information under international agreements, consider two states, or groups of states, that perceive themselves to be in a PD relationship with strong expectations of iteration. Both sides see it as advantageous to cooperate and to manifest their cooperation in an agreement.<sup>47</sup> Consider first the kinds of information these states will seek at the outset, as they negotiate their agreement, then the information they will seek on an ongoing basis under the agreement.

---

44. The breakup of the Soviet Union and Yugoslavia shows that international interactions are not indefinite. Dramatic events like those, however, are neither frequent nor readily predictable.

45. See Abbott, *supra* note 1, at 363-64. Horizontal iteration can be intentionally increased through strategies of issue linkage. For example, the Helsinki Final Act essentially linked acceptance of human rights norms by the Soviet Union to implicit recognition of the postwar political situation in Europe.

46. Models that fail to take account of factors like time and issue linkage have overlooked the importance of reputation. For a discussion of temporal considerations in connection with a famous model, see Bendor & Hammond, *supra* note 8, at 306.

47. States may choose to make such agreements less formal than traditional treaties. See Charles Lipson, *Why Are Some International Agreements Informal*, 45 INT'L ORG. 495 (1991).

At the outset, the crucial facts relate to the nature of the interaction: the strategic structure of the relationship (whether it is actually a PD or some other game) and the conditions under which the game will be played. Many PD analyses assume that all players have complete information on the structure of the game.<sup>48</sup> In international politics, however, the parties are likely to suffer from "prospective uncertainty"<sup>49</sup> as to all the basic determinants of game structure: (a) the courses of action available to them and the results of different combinations of choices; (b) the preferences of the other side, which determine the payoff structure; and (c) the conditions of play.

(a). The parties to an agreement will typically face options more complex than a binary C/D choice.<sup>50</sup> In some circumstances, this can make the situation difficult for the players to understand, as well as for observers to model.<sup>51</sup> The factual outcomes associated with pairs of actions are, in the real world, complex matters of probability influenced by a variety of factors. The parties are likely to expend considerable effort analyzing these matters, but they are largely exogenous to the relationship itself.

(b). While the PD and similar games are usually presented as symmetrical, at least in ordinal terms, the two parties to an interaction may not have the same preferences; even if Party I perceives a relationship to be a PD, Party II may be playing a different game.<sup>52</sup> Its preferences may be more conflictual than Party I realizes. For example, Party II's true goal may be to minimize the payoffs of the other players, in which case it will try to defect regardless of iteration.<sup>53</sup> Party II may also have less conflictual preferences than Party I believes.<sup>54</sup> In this situation, Party I

---

48. See *supra* text accompanying note 8.

49. Bendor & Hammond, *supra* note 8, at 303.

50. States typically have many gradations of choice, as shown by the fact that most disputes over treaty compliance involve neither obvious compliance nor outright breach, but something in between. See DOWNS & ROCKE, *supra* note 8, at 7.

51. According to Snidal, even a large number of options will not fundamentally change the game so long as they occupy a single policy axis, from free trade to protection. If the choices span different axes, however, the nature of the game can become quite unclear. See Snidal, *supra* note 5, at 927-28.

52. Goodin, for example, suggests that the U.S. and the Soviet Union had different preference orderings during much of the Cold War. While each periodically favored cooperation, their cooperative periods rarely coincided. See Robert E. Goodin, *Research Note: Mood Matching and Arms Control*, 32 INT'L STUD. Q. 473 (1988).

53. Similarly, Party II could have the preferences of the Deadlock game, which would make entering an agreement unlikely. See Abbott, *supra* note 1, at 357-58. Some analysts argue that Deadlock preferences underlie many arms races. See DOWNS & ROCKE, *supra* note 8, at 75; George W. Downs et al., *Arms Races and Cooperation*, 38 WORLD POL. 118, 120-23 (1985). Risse-Kappen argues that Brezhnev and Gromyko were playing Deadlock in the early 1980s with regard to intermediate-range nuclear forces. See Thomas Risse-Kappen, *Did "Peace Through Strength" End the Cold War? Lessons from INF*, 16 INT'L SECURITY, Summer 1991, at 162.

54. Plous argues that the United States and the Soviet Union were for decades unable to ease their military competition because each, probably erroneously, saw itself as playing a cooperative game and the other as playing a competitive game. See S. Plous, *Perceptual Illusions and Military Realities: The Nuclear Arms Race*, 29 J. CONFLICT

could unnecessarily weaken the cooperative relationship by ambitious preparations for defensive defection. Here, however, Party II would have a strong incentive to make its true preferences known.

Each state will also seek information as to the intensity of the other's preferences.<sup>55</sup> If, for example, Party I knew that Party II perceived the result of unilateral cooperation to be a national disaster, or perceived the reward for unilateral defection to be extremely high, it would be particularly cautious about entering a cooperative relationship at all; it would at least seek to include appropriate provisions in the agreement.<sup>56</sup> If Party I were counting on its ability to respond to Party II's conduct under the agreement with actions in a different issue area, it would also want to know how Party II values gains and losses as between those areas. Again, if Party II strongly favored cooperation, it would have reason to make its preference known.

Unfortunately, information about preferences can be extremely difficult to obtain and to convey.<sup>57</sup> In a complex collective entity like a state, full sets of cardinal or interval preferences may never be clearly defined. Even if they exist, they will be difficult to document, since many sources of information will be secret, self-serving or distorted for political consumption. States with conflictual preferences, moreover, will have strong incentives to misrepresent them.<sup>58</sup> States will often have to make do with incomplete knowledge of preferences, relying on reputation, recent experience and their own understanding of the situation as proxies.<sup>59</sup>

(C). The most important condition of play, at least if the relationship is PD, is the degree of iteration. States negotiating an agreement will attempt to divine the future course of their relationship in the issue area, as well as the likely opportunities for issue linkage. Even if the relationship appears indefinitely iterated, though, the success of conditional cooperation also turns on the extent to which states are influenced by expectations of future responses. Each party will thus seek

---

RESOL. 363, 382 (1985). Such a situation can become self-sustaining if each side interprets the continued absence of agreement as confirming the other's intransigence. *Id.*

55. Intensity is irrelevant in the standard PD. The dominant strategy of defection depends only on the ordinal structure of preferences.

56. Strong assurance and verification devices, respectively, would be among the appropriate provisions.

57. For an absorbing account of the Cuban missile crisis as a game of incomplete information, showing how difficult it can be to determine the preferences of another state, or even to credibly reveal one's own preferences, see R. Harrison Wagner, *Uncertainty, Rational Learning, and Bargaining in the Cuban Missile Crisis*, in ORDESHOOK, *supra* note 8, at 177.

58. See Abbott, *supra* note 1, at 402-04.

59. A state may demand that others take steps, like joining an international regime, to manifest their commitment to certain preferences as a condition of negotiating further agreements. The United States, for example, did not think seriously about negotiating a free trade agreement with Mexico until the latter joined GATT. Like other representations of preferences, however, such steps may be unreliable.

information on “the shadow of the future”.<sup>60</sup> the other party’s time horizon—how far into the future its government is willing and able to look—and the degree to which it discounts future gains and losses compared to current ones.<sup>61</sup> These characteristics, however, are subject to the same informational problems as other types of preferences.

Consider now the types of information the two sides will be concerned with on an ongoing basis, the issues they may find it necessary to address in their agreement. Clearly, each party will seek current information as to the structure of the interaction, both because its original information may have been inaccurate and because strategic situations are subject to change. Options, outcomes, conditions of play and all forms of preferences can evolve over time. They may also shift abruptly with changing interest group pressures and popular perceptions, changes in government—from Carter to Reagan, or Brezhnev to Gorbachev<sup>62</sup>—technological advances, and other events.<sup>63</sup> Unfortunately, such information is at least as problematic on an ongoing basis as it was initially.

Once an agreement is actually in operation, however, information as to the conduct of the parties under the agreement also becomes relevant, and can be of crucial importance. As discussed earlier, the capacity to recognize other parties’ actions is essential to the strategy of conditional cooperation and to one’s ability to avoid the “sucker’s payoff” by changing an earlier move. In many situations, a party can achieve acceptable payoffs even without understanding the structure of an interaction so long as it has accurate information on the actions of others.<sup>64</sup> Information about conduct, especially patterns of conduct, can also support inferences as to other issues, such as current preferences. For all of these reasons, each party will undoubtedly seek to gather information on the conduct of others, focusing particularly on indications of defection or impending defection. At the same time, each party may also seek to provide others with information as to its own conduct: helping others to recognize one’s cooperative and non-cooperative actions can make conditional cooperation all the more effective, and may help protect the agreement from the risk of defensive defection.

Traditional analyses of the PD have assumed either that players have complete information about prior moves or, in the case of one-shot games where the players move simultaneously and cannot communicate, that such information is irrelevant or unattainable. Neither assumption is realistic. There is often considerable “retrospective uncertainty” as to

---

60. See AXELROD, *supra* note 22.

61. See Abbott, *supra* note 1, at 364.

62. For an argument that only the latter change in leadership made possible the INF Treaty, and by implication other recent arms control agreements, see generally Risse-Kappen, *supra* note 53.

63. The greatest changes may result from technological developments that permit actions capable of ending or substantially altering the game, such as a surprise attack or *blitzkrieg* invasion.

64. See Wagner, *supra* note 21, at 335.

the actions of other parties to an agreement, a condition exacerbated by the existence of incentives to mislead.<sup>65</sup> Yet states can normally improve the quality of the information available, often far more easily than with subjective matters like preferences.<sup>66</sup> Since information about conduct is endogenous to a relationship, states can jointly facilitate its provision by including appropriate mechanisms in their agreements.

## II. Verification and Assurance

Two basic strategies for the production of information, each with its characteristic procedures, can be derived from the twin centrifugal incentives of the PD game: the offensive and defensive motives to defect.

### A. Offensive Defection

The offensive incentive is an incentive of temptation. It leads parties to cheat or shirk on their commitments in ways they hope will not be observed, rather than to withdraw from those commitments altogether, since defection is only profitable when it is unilateral. Similarly, the offensive incentive leads defecting parties to conceal or disguise their actions and to mislead others about them. At least in iterated and second-chance games, where effective responses are available, offensive defection is largely a problem of inadequate information.

States negotiating an agreement in a PD setting will almost always seek ways to minimize offensive defection by others. These efforts will take on special urgency, however, if states suspect that others harbor, or may come to harbor, predatory, difference-maximizing or similarly conflictual preferences, or that others value the payoff for unilateral defection very highly relative to that for mutual cooperation. Concern over offensive defection will also be strong when the overall relationship between the parties is one of suspicion—based on reputation or prior experience—or unfamiliarity.

Iteration tends to weaken the offensive incentive, but offensive defection will remain a serious concern if states believe that others have, or may come to have, short time horizons or high discount rates for future payoffs.<sup>67</sup> Concern for offensive defection will also dominate if states fear the development of technological advances—the breakout of an advanced weapons system, for example—or innovative policies capable of producing one-time gains large enough to offset the likely future costs or, in the extreme case, to end the game.

---

65. See Abbott, *supra* note 1, at 366-67.

66. Interpreting such information and characterizing conduct as compliance or non-compliance, however, represent much more complex and subjective matters.

67. A state may be unwilling to value the future heavily because of limited governmental terms of office and electoral considerations; serious social or economic problems, or anticipated changes of national status.

## B. Verification

When concern over offensive defection by others exists in a relationship, the parties will typically rely on the strategy of verification.<sup>68</sup> In verification, each party uses its own efforts and resources to monitor<sup>69</sup> or gather information about the actions of the others. Verification is a negative strategy. Suspicious states attempt to police the conduct of others with the aim of preventing them from taking the kinds of action they are thought to prefer. Since the other parties are expected to cheat, they cannot be trusted to produce accurate information. Thus each state keeps the information-gathering function wholly in its own hands.<sup>70</sup> Even then, the others may be expected to interfere, subtly or blatantly. When international agreements impose symmetrical obligations, each party is typically both verifier and verifiee, though some may seek information more aggressively than others.

Certain characteristic mechanisms and procedures are associated with the strategy of verification. To introduce some of these, and to elucidate the strategy further, consider the verification techniques that might be relied on in connection with the following hypothetical agreements.<sup>71</sup>

1. Amerigo and Nipporea have agreed in principle that their fishing fleets will no longer use large drift nets, in order to protect ocean mammals.

Under a pure verification approach, each party would institute procedures designed to monitor the activities of the other's fishing vessels. The techniques used would be relatively straightforward, since the relevant obligation relates to readily observable physical activities. Each

---

68. Devices intended to minimize the risk of defection by others under an agreement will typically limit one's own ability to defect as well. Such reciprocity is normally a necessary element of the bargain. Presumably, however, each state will attempt to negotiate devices that restrict its likely activities less than those of others. For a similar problem in connection with confidence-building measures, see *infra* note 203.

69. In many discussions of arms control verification, the gathering of information is referred to as "monitoring", or a similar term, while the analysis of that information—including the often subjective process of determining whether it reveals a violation—is referred to as "verification." See, e.g., Robert J. Einhorn, *Treaty Compliance*, 45 *FOREIGN POL'Y*, Winter 1981-82, at 30; Patricia B. McFate & Sidney N. Graybeal, *The Price for Effective Verification in an Era of Expanding Arms Control*, 500 *ANNALS* 73 (1988). In this article, I use the terms interchangeably, to refer to the information-gathering phase. The substantive distinction is important, however. Verification and assurance devices will, or should, be designed to produce information suitable for use in the evaluation phase; they may also be manipulated to influence the outcome of that phase.

70. As noted in the introduction, states may also entrust this function to an international organization or to some other party thought to have superior capabilities. See *supra* note 17. For a further discussion of international verification organizations, see *infra* notes 109-20 and accompanying text.

71. These hypotheticals are intended as abstract examples. I do not warrant that they are accurate representations of actual arrangements either politically or technically.

party would presumably utilize its own vessels and aircraft to observe the other's fleet while it is engaged in fishing operations. This kind of observation would not require any special authorization so long as the fleets were operating on the high seas.

Fishing fleets, however, operate across wide stretches of ocean, and their nets are not always in use. Comprehensive observations from ordinary commercial vessels and aircraft thus might be infeasible, while special flights dedicated to monitoring would be quite costly. Unfortunately, fairly comprehensive observations would be needed in this case, since the aim is to verify a negative. Observation satellites might supply a technological answer, but only a limited number of states possess them, and they might well be dedicated to more urgent functions. If membership in the drift net agreement later increased, some of these coverage problems could be ameliorated through arrangements for the pooling of information.

Without special authorization, each party could also examine publicly available documents on fishing procedures published by the other party. These might not be very revealing, however, especially if either side intended secretly to violate the agreement. Private business travelers are often able to observe conduct under agreements; here, however, such observations are not likely to be satisfactory.

In their agreement, then, the parties might find it necessary to arrange for monitoring procedures that would be impermissible without special authorization. The agreement could, for example, authorize agents of each party to observe the unloading of catches in the ports of the other party, or in third country ports if the local authorities also agreed; to send observers on selected voyages of the other party's fleet; or even to board vessels at sea in unannounced spot checks. Procedures like these would almost certainly have to be authorized on a reciprocal basis and might be very costly. For both reasons, the parties might well agree to limit the scope and frequency of direct observations, even though this would reduce their probative value.

2. Amerigo and Canamex have agreed to reduce emissions of sulfur dioxide (SO<sub>2</sub>) suspected of causing cross-border acid rain.

Under a pure verification approach, the parties would once again institute procedures designed to monitor the activities subject to the agreement, SO<sub>2</sub> emissions. Physical observation would be more difficult than in the drift nets case, however, since emissions are produced within each party's national territory. Without special authorization, each party could only collect air samples within its own territory, perhaps in areas with prior cross-border emissions problems, or at border points close to the other party's industrial concentrations. More indirectly, each could take soil or water samples in areas subject to acidification. Sampling on one's own territory might be inadequate, however, without enough additional information to establish causation: information ranging from wind patterns to production schedules and fuel usage in the other

party's factories. In addition, since the agreement restricts emissions rather than prohibiting them altogether, fairly detailed and comprehensive data would be necessary to determine compliance.

Each party might also attempt to monitor written data on SO<sub>2</sub> emissions from factories, power plants, and similar sources within the other party's territory, but this approach would be futile unless both governments collected the relevant data from private firms and placed it on the public record. Some information might be available through the press and non-governmental organizations, but it would probably be incomplete, if not anecdotal. Internal corporate data would in all likelihood be confidential and protected by law.

In the end, each state might again consider requesting authorization to send specialized agents into the other's territory. Such agents could monitor factory emissions on a spot or ongoing basis or observe the installation of emissions-control equipment under conditions specified in the agreement. Any significant program of inspections, though, would be costly and might cause considerable inconvenience. As a result, direct observations would almost certainly have to be limited in scope and frequency. Legal questions relating to private property and confidential corporate information might well arise. With the burden of reciprocity, both states might find this kind of authority difficult to grant.

These simple examples suggest several important points about the strategy of verification:

First, whether verification will be successful depends strongly on the nature of the norms and conduct at issue. Both hypothetical cases suggest, for example, that it will typically be easier to verify compliance with positive obligations than with prohibitions. Thus, if the Amerigo-Nipporea agreement called for both parties to destroy their stocks of drift nets, or to change the configuration of their fishing vessels so that drift nets could not be used, compliance could be directly observed. As it is, though, compliance can only be fully determined by observing all fishing vessels at all times.

The cases also suggest that it will be easier to obtain information on physical activities than on intangibles, and on government actions than on private actions (at least in liberal states with strong private property regimes).

Second, while verification devices can be quite simple, they can also become elaborate and costly if technical activities like satellite observation, large-scale inspections and the like are necessary.

Third, effective verification procedures will often require intrusive measures: access to foreign territory, to the operations of government agencies and private firms, and to bodies of evidence that would otherwise be confidential. While states can monitor many kinds of conduct without any special arrangements, such intrusive approaches are both infeasible and unlawful without special authorization. At this level, the verification strategy cannot work without the acquiescence, or even the

participation, of the verifiee.<sup>72</sup>

Fourth, because they are so intrusive, elaborate verification schemes will often be resisted, for reasons of sovereign principle, out of concern that others will use them to gather valuable information on unrelated matters, or simply to avoid interference with legitimate activities. At the least, states are likely to seek to limit the scope of intrusive verification procedures in explicit and fairly detailed fashion, even though this may limit their usefulness.

Finally, for all of the foregoing reasons, the verification strategy alone may be incapable of producing enough reliable information to make cooperation possible, at least at an acceptable cost.

### C. Defensive Defection and the Stag Hunt Game

Consider now the defensive incentive to defect in the PD game. This is an incentive of fear. It motivates a state to abandon cooperative arrangements, not in pursuit of gain, but in an effort to avoid the losses that it believes others will inflict upon it. Defensive defection can take the form of non-compliance or of complete withdrawal from an agreement seen as unacceptably risky. Non-complying conduct may be concealed, like offensive defection, or open, as in the case of countermeasures designed to respond to defection by others. The defensive incentive to defect—like the self-defense rationale more generally—may lead a state to act preemptively, especially when it fears that it would be unable to recover fully from a “sucker’s payoff.”

Defensive defection can figure in the calculations of states entering an agreement in two ways. First, when the parties are concerned with the possibility of offensive defection by others, each may seek to guarantee its own ability to defect defensively. General international law may provide sufficient opportunities,<sup>73</sup> but states may include specific authority for countermeasures or withdrawal in their agreement. Most major arms control agreements, for example, authorize any party to withdraw “if it decides that extraordinary events related to the subject matter of [the] treaty have jeopardized its supreme interests.”<sup>74</sup> States will rely on verification mechanisms for timely warning.

Second, and more relevant for present purposes, each state may seek ways to avoid defensive defection by others. This problem is best

---

72. In effect, acquiescence or participation functions as an assurance device, a way to convey the information that one is complying with one’s obligations. Most agreed arrangements designed to facilitate monitoring are to some extent a blend of the strategies of verification and assurance. See *infra* pp. 26-27, 29-30.

73. See, e.g., RESTATEMENT (THIRD) OF FOREIGN RELATIONS LAW OF THE UNITED STATES § 905 (1980) (state injured by violation authorized to take proportional measures of self-help); Vienna Convention on the Law of Treaties, May 23, 1969, art. 60, S. Executive Doc. L, 92d Cong., 1st Sess. 27 (1971) (material breach entitles other parties to invoke breach as ground for termination or suspension of treaty).

74. See, e.g., ABM, *infra* note 98, art. XV ¶ 2, 23 U.S.T. at 3446.

illustrated by the game known as Stag Hunt (SH).<sup>75</sup> In the SH story, a band of primitive hunters is attempting to capture a stag. If all of them stay with the hunt, they can encircle the stag; if any of them abandons the hunt, the stag will escape. Stag is the hunters' favorite food, so all prefer a successful hunt. There are, however, other good things to eat in the forest. One or more of the hunters could observe a rabbit and be tempted to break off from the hunt to capture it. These hunters will eat lightly if they do so, but they will cause the stag hunt to fail, and those who continue after the stag will not eat at all. Under these circumstances, all would prefer the hunt to end without a hare out of place.<sup>76</sup> The SH matrix is shown in Figure 4.

The structure of preferences in SH is highly conducive to cooperation, even in a one-shot game. Mutual cooperation (a stag dinner) is the preferred outcome for all.<sup>77</sup> Mutual defection (rabbit for all) produces a lower payoff. The striking feature of the game, however, is that unilateral defection also produces a lower payoff: even the hunters that eat rabbit would prefer stag.<sup>78</sup> Unlike PD, then, the players in a SH game have no offensive motive to defect.<sup>79</sup> They do, however, have a defensive motive to defect, to avoid fruitless pursuit of the stag once its capture has become impossible.<sup>80</sup> The typical SH situation is one in which two or more states or other actors seek an outcome that requires unanimous participation.<sup>81</sup> The SH game has been used to analyze the pro-

---

75. The game is drawn from a parable told by Rousseau. JEAN-JACQUES ROUSSEAU, DISCOURSE: ON THE ORIGINS AND FOUNDATIONS AMONG MEN (1975), reprinted in THE FIRST AND SECOND DISCOURSES (Roger D. Masters ed. & Roger D. Masters & Judith R. Masters trans., 1964). Rousseau's original story is ambiguous in important respects, and can be interpreted as more or less conducive to cooperation. See Abbott, *supra* note 1, at 368-69.

76. Lines like this are a tradition in discussions of Stag Hunt, though it should be clear from the text that the best ones have already been taken. See, e.g., RUSSELL HARDIN, COLLECTIVE ACTION 169 (1982) ("hare today, gone tomorrow").

77. As with the discussion of PD, this section assumes only two players, avoiding some of the complexities of multiple-player games.

78. Because the matrix in Figure 4 shows the ordinal preferences 1-4, it seems to suggest that a hunter would value a rabbit more highly if she alone were eating one than if all the hunters were eating rabbit. This might be true if, for example, there were not enough rabbits to go around. This aspect of the game is not important to the current analysis. See Duncan Snidal, A Problem of Assurance, The ROC and Taiwanese Investment in the PRC (Dec. 1990) (unpublished manuscript, on file with Cornell Int'l Law Journal).

79. There is only a single difference between the SH matrix and the PD matrix: each player's first- and second-highest preferences are reversed.

80. Since all SH players prefer to cooperate so long as all the others do, and prefer to defect if any others have, there is no dominant strategy in the SH game, as there is in the one-shot PD game. All choices are interdependent.

81. Even arms control could be a SH game. See, e.g., DOWNS & ROCKE, *supra* note 8, at 76. Most arms control agreements are signed when rival states reach military parity and weapons development is static. See Joseph Kruzell, *From Rush-Bagot to START: The Lessons of Arms Control*, 30 ORBIS 193, 197 (1986) (conditions like these are typical; arms control legitimizes status quo); Raymond L. Garthoff, *SALT I: An Evaluation*, 31 WORLD POL. 1 (1978) (observing such conditions at the signing of SALT I, which froze status quo). In these conditions, states might prefer the stability

Figure 4

		Player II	
		C	D
Player I	C	4, 4	1, 3
	D	3, 1	2, 2

C = Cooperation                      D = Defection

duction of certain types of collective goods,<sup>82</sup> as well as other situations in international politics where widespread or unanimous cooperation is necessary.<sup>83</sup>

In the standard account of SH, like that of PD, it is assumed that all players know the structure of the game<sup>84</sup> but are ignorant of each other's actions.<sup>85</sup> Under these conditions, each player should be able to observe that no other player has any offensive motive to defect. Without actual knowledge of the other players' conduct, though, some might still suspect that others were not cooperating, perhaps through inadvertence, through a misunderstanding of the options, or in response to a sudden temptation. If fear of the "sucker's payoff" were strong, these suspicions could induce such states to defect defensively, even preemptively. In turn, fear of such defections could induce other defensive defections, destroying an agreement that all wished to maintain and with which all were in fact complying.<sup>86</sup>

---

of the status quo to all other options, even unilateral arms development, so long as all shared the same preferences.

82. See, e.g., Michael Taylor & Hugh Ward, *Chickens, Whales and Lumpy Goods*, 30 POL. STUD. 350 (1982); Carlisle F. Runge, *Institutions and the Free Rider: The Assurance Problem in Collective Action*, 46 J. POL. 154 (1984); Hugh Ward, *Taking Risks to Gain Reassurance in Public Goods Games*, 33 J. CONFLICT RESOL. 274 (1984).

83. See Abbott, *supra* note 1, at 368-71; Lisa L. Martin, *Institutions and Cooperation: Sanctions During the Falklands Islands Conflict*, INT'L SECURITY, Spring 1992, at 143, 145; Snidal, *supra* note 78.

84. Strictly speaking, each player must also know that all the other players know that all know the structure of the game, and so on, in infinite regression. In theory, this "common knowledge" problem defies analysis; in international interactions, even though states can realistically take into account only a very few levels of the problem, it can still pose practical problems. See Snidal, *supra* note 78.

85. The same problems would exist, though in less extreme form, if knowledge of the actions of other players were incomplete, rather than non-existent.

86. If the players in a SH game also had less than full confidence in their understanding of the structure of the game, the situation would be much different: others might have offensive motives to defect; the game might not be SH at all. For further discussion, see *infra* p. 26.

Although inadvertence, misunderstanding and temptation should not be ignored, the real danger in the SH game is erroneous defensive defection, and this is wholly a problem of inadequate information. What is needed to ensure continued cooperation is a way for each player to communicate its ongoing commitment to the cooperative result in a credible fashion. States entering an international agreement in a SH setting can best deal with the inherent instability of the game by establishing mechanisms to ensure the production of information revealing their own preferences and compliance, either independently or in the agreement.<sup>87</sup>

The problem of avoiding erroneous defensive defection also arises in PD games. In theory, at least, it should be of significant concern to the negotiators of international agreements in a wide variety of settings.<sup>88</sup> This concern is likely to be more prominent, however, when the parties are less preoccupied with the risk of offensive defection: when there exists, to a reasonable degree, a relationship of trust or a belief that the other parties prefer cooperation, a high level of confidence in the effects of iteration, and a sense that policies and technologies are relatively stable. Indeed, if all conditions are right, a situation that might be called "perfect iteration," an iterated PD game can become the virtual equivalent of SH.<sup>89</sup> In circumstances like these, erroneous defensive defection will be the main threat to continued cooperation.

#### D. Assurance

When a desire to avoid defensive defection by others figures strongly in a relationship—most clearly in SH situations, but in most PDs as well—states will typically rely on the strategy of assurance. In assurance, each party to an agreement expends its own efforts and resources to assemble information about its own activities and to convey that information to the other parties. Assurance is a positive strategy. It is used when each party believes that the others, like itself, wish to maintain cooperation under the agreement, but seek to avoid being taken advantage of, by accident or design. Assurance devices help each party to give others confidence in its own cooperation, with the aim of permitting all parties to follow the cooperative course of action they prefer.<sup>90</sup>

---

87. The inclusion of such mechanisms in the text of an agreement will help to assure all parties that the problem has been addressed at the outset.

88. Even a state that harbors hopes of offensive defection will have reason to support mechanisms that forestall defensive defection by others. The potential defector, after all, must maintain the agreement in force if it is to reap the full benefit of its actions. The last thing it wishes to see is the premature collapse of the agreement based on erroneous perceptions.

89. The players must perceive that the interaction will be repeated indefinitely, and the players' time horizons and discount rates, expectations of recognition and response, and preference intensities must be such that the long-run benefits of future cooperation are seen to outweigh any potential short-run gains from defection. The equivalence to SH is derived algebraically in Snidal, *supra* note 78.

90. As noted earlier, however, the assurance strategy may also be manipulated by states planning to defect. See *supra* note 56.

Since maintaining the confidence of others is the primary goal, no party wishes to rely on the chance that others will come by sufficient information on their own. Assurance therefore places the process of assembling and conveying information about each party in the hands of that party itself. Assurance is, however, a more interactive strategy than verification. For assurance to be effective, the assuring state must first recognize that others need assurance; it must then provide them with information in a manner certain to reach them and in a form credible to them. When symmetrical obligations are imposed, each party will be both assurer and assuree, although some—including those whose actions are otherwise less transparent or who have developed a reputation for responding rashly to temptation—may find it necessary to provide more information than others.

Like verification, the assurance strategy entails the use of certain characteristic mechanisms and procedures. To introduce these, consider how the two hypothetical agreements discussed above<sup>91</sup> might be structured if the parties relied solely on assurance devices.

1. The agreement between Amerigo and Nipporea to restrain the use of drift nets.

Under a pure assurance approach, each party would institute procedures designed to gather and transmit to the other credible information demonstrating its own compliance. If the parties shared a high degree of trust, simple verbal assurances might suffice. Otherwise, more concrete procedures would be necessary. Direct evidence that particular vessels were not using drift nets, though, might be difficult to provide.

As a form of indirect evidence, each party might publicly promulgate legal restrictions on the use of drift nets by private vessel operators in its fleet. These restrictions could be backed up with visible enforcement activities. Each party might also require vessel operators to file certifications that drift nets had not been used, along with inventories of the species of marine life found in each catch; these documents could be made public or supplied directly to the other parties. In order to render these measures credible, however, the parties might have to establish procedures to corroborate the private reports. Each party could, for example, require that its vessels be unloaded before government agents, who would publicly certify the content of catches.

Taking another tack, each party might agree to require the physical conversion of the fishing vessels in its fleet so that they were manifestly unable to use drift nets. Such actions would demonstrate compliance in a manner directly observable by the other side. Each party could further support its undertaking with a program of periodic inspections and public reports, and could even invite agents of the other party to participate in such inspections.

---

91. See *supra* text accompanying note 71.

## 2. The agreement between Amerigo and Canamex to restrain SO<sub>2</sub> emissions

As in the drift net case, each party could convey indirect assurances of compliance by adopting formal emissions requirements for factories and similar pollution sources within its territory and implementing visible oversight and enforcement processes. Each government might collect emissions data, from its own agencies' observations or from corporate reports, and supply the data to the other party.

More concrete, and thus costly, assurance devices might rely on technology. For example, on the output side, each party could require major factories to install devices that would automatically measure SO<sub>2</sub> emissions. Such measurements could be included in periodic reports or relayed directly to the other party. Alternatively, each could require the installation of approved devices to remove SO<sub>2</sub> from emissions, with public inspection of compliance, including observation by the other party's agents. On the input side, each party could require installation of furnaces capable of burning only low-sulfur fuels, again with public certification of compliance.<sup>92</sup> Like the conversion of fishing vessels, this approach would make compliance directly visible to the other side.

Many of the points previously suggested by these examples apply to assurance techniques as well as verification:

First, the success of the assurance strategy depends on the nature of the obligations in question and the types of information needed. In the two hypothetical cases, for instance, each government has the doubly difficult task of providing assurance of a negative that involves private conduct.

Second, while assurance procedures can be extremely simple, they can also become complex and costly, as with the technological solutions discussed above. Assurance techniques tend to be especially complex when individuals and firms within each party, or the parties themselves, may have incentives to mislead, for each action taken to assure must often be followed by a second action to establish the credibility of the first.

Third, assurance devices can often be made more credible if the assuree is invited to participate in them, just as verification can be made more effective if the state under observation permits greater access by foreign monitors. Such participation, of course, leads to the kinds of intrusion that assurance otherwise avoids.

Finally, like verification, assurance alone may be unable to provide enough information for cooperation to proceed.

## E. Verification and Assurance as Complementary Strategies

In virtually all PD relationships, states will have and be perceived to have

---

92. To repeat, these are intended as abstract examples of verification and assurance techniques, not as discussions of practical solutions.

both offensive and defensive motives to defect.<sup>93</sup> One or the other motive may appear stronger; one may even appear dominant; but both will be present. When states enter into international agreements in these situations, one would expect them to provide for both verification and assurance procedures, or some hybrid of the two strategies, with one or the other more prominent depending on perceptions of the situation.

For a variety of reasons—the absence of any central authority to enforce compliance, the pervasive uncertainty of international relations, the possibility of highly conflictual preference orderings, and the climate of suspicion that surrounds many international relationships—verification is likely to be the dominant approach in most PD situations. Since states with an offensive incentive to defect have a concomitant motive to mislead, relying solely or even primarily on assurance devices would be foolhardy. Indeed, excessive reliance on pure assurance could increase the likelihood of offensive defection, a version of the problem of moral hazard.

For many of the same reasons, verification techniques are likely to be used, if only as backups, even when the underlying relationship appears to reflect SH preferences or highly cooperative conditions approaching the perfectly iterated PD game. One may have read the strategic situation incorrectly, or conditions could subsequently change. Even aside from these problems, states have reason to seek all the information they can obtain at a reasonable cost. When entering an agreement in a SH situation, for example, it will often be reasonable for each party to expend some resources attempting to verify that others are not defecting out of inadvertence or misunderstanding, even if each is simultaneously attempting to assure the others of its continued compliance. Assurance alone may simply be unable to provide sufficient reliable information. By the same token, states seeking to convey assurances may find some foreign monitoring desirable as a way to channel information; they may even invite some participation by other states, incorporating a form of verification into the assurance procedures themselves.

After all this, one might ask why states would ever rely on assurance devices: if states are likely to engage in verification even in highly cooperative situations, why not rely solely on that strategy? There are, however, powerful reasons to utilize assurance whenever a problem of premature or erroneous defensive defection may arise.

The most basic reason is that verification has its limits; it will often fail to produce adequate information on its own. Well-designed assurance devices, moreover, can often provide more complete and accurate information than external monitoring on certain subjects, since the responsibility for providing the information is placed on the state with

---

93. "Perfectly iterated" PD games, similar to SH games, may be an exception. See *supra* note 89.

the greatest knowledge of and access to the activities in question. For much the same reason, assurance may be able to provide a given class of information at considerably lower cost. It is for these reasons that states engaged in verification request the participation of the verifiee, incorporating a form of assurance.

In addition, the assurance strategy may avoid certain political difficulties common to verification. As discussed above, effective verification techniques are often highly intrusive and can be over-inclusive, posing a risk of opportunism.<sup>94</sup> The use of assurance devices can ameliorate this difficulty, though the characteristic problem of assurance, credibility, means that it will rarely be a complete solution.

More positively, assurance tends to actively commit states to a cooperative relationship. Each state's informational activities are oriented toward the goals of promoting cooperation and keeping the agreement intact. Domestic bureaucracies and procedures are organized to further this goal, creating vested interests in the success of the relationship. Verification, in contrast, leads to no such mobilization of commitment, but instead leads states to organize their technical procedures, bureaucratic organization and political attitudes around the assumption that other parties to the agreement are likely to cheat. From the perspective of both assurer and assuree, the use of at least some assurance devices may help relationships to develop along more cooperative lines.

In the end, the ideal solution for international agreements in situations where information is important may be a hybrid of assurance and verification techniques. Such a combination, if carefully designed, could take advantage of the strengths of each strategy while offsetting its weaknesses. The drift net and SO<sub>2</sub> hypotheticals, for example, both suggested arrangements in which the two sides would participate jointly in the production of information, blending verification and assurance.

#### F. Verification and Assurance in International Agreements

As noted earlier, parties to international agreements can utilize many verification and assurance techniques even if no formal arrangements are made *ex ante*. Private firms, for example, can easily monitor foreign compliance with obligations to lower tariffs, receive foreign investments, protect intellectual property rights, and the like; as a result, international commercial agreements rarely spell out explicit verification procedures. Government officials can monitor foreign compliance with many other kinds of agreements—boundary settlements, air service agreements, extradition treaties, and a host of others—without special arrangements. Similarly, governments can convey assurances of compliance in a variety of ways, both verbally, through diplomatic channels or more informally, and non-verbally, through acts designed to reveal a

---

94. In arms control negotiations, for example, the Soviet Union long resisted external verification as a form of espionage. See Allan S. Krass, *Verification and Trust in Arms Control*, 22 J. PEACE RESOL. 285 (1985).

course of conduct, whether or not their agreement specifies any particular procedure. To understand fully the workings of international agreements, then, it will often be necessary to look beyond their explicit provisions: the informational mechanisms that help maintain cooperation may operate wholly outside these institutional bounds.

It is also important, however, to ask why states sometimes find it necessary or desirable to spell out explicit arrangements for the production of information *ex ante*, in their agreements. The preceding discussion of verification and assurance has suggested several rationales for the inclusion of such provisions and for the forms they are likely to take.

First, the discussion has suggested three principal rationales for explicit verification provisions:

1. Most fundamentally, states will tend to spell out explicit verification procedures when the modes of observation they deem necessary to monitor potential defection cannot be carried out without special authorization. Thus, referring again to the hypothetical cases discussed earlier, if international law, or simply the political realities of the situation, would make it infeasible to board the fishing vessels of another state, dispatch air quality inspectors into its territory, or send monitoring aircraft over its territorial sea, the states negotiating an agreement on drift nets or SO<sub>2</sub> emissions might find it necessary to expressly legitimate such activities in order to achieve cooperation on the substantive issues.

2. States entering an international agreement may plan to rely on a third party for the gathering of information. All parties may agree that an international organization, or even a non-governmental organization, should monitor conduct under the agreement. Alternatively, certain parties may agree that one of them will perform the monitoring function on behalf of all. In cases like these, states will typically have to grant the third party the authority to act as monitor, as well as authorizing specific modes of observation. Agency provisions like these will often have to clarify a variety of incidental issues, such as the manner in which information will be shared among the parties.

3. Since the relatively intrusive modes of observation likely to require specific authorization in international agreements have the potential to gather significant amounts of sensitive collateral information, as well as to interfere with legitimate activities, verification provisions are likely to place limits on the monitoring activities they authorize, and conceivably on independent modes of monitoring as well.

The discussion in this section has also suggested several important rationales for explicit assurance provisions, many of which derive from the interactive nature of the assurance strategy:

1. Most generally, by agreeing to an explicit assurance procedure, states can manifest a clear desire for cooperation and can commit themselves to give specific assurances in the future. On both counts, such an *ex ante* commitment should increase the confidence of other states,

making them more willing to enter into an agreement. For this reason, an explicit assurance provision may appear desirable.

2. More specifically, an explicit provision allows all parties to an agreement to spell out in advance the forms of assurance that will be acceptable. In their capacity as recipients of assurances, the parties may only be willing to accept information of a certain scope, gathered and presented in particular ways, as credible and probative. In their capacity as assurers, on the other hand, the parties may only be willing or able to provide relatively limited classes of information, in limited forms. Negotiation of an explicit assurance provision at the outset allows the parties to harmonize their needs and capabilities before situations engendering concern have arisen, avoiding disappointment and the escalation of suspicion later on.

3. An explicit assurance procedure can increase the effectiveness of the interactive communications essential to the success of the strategy. From the perspective of a potential assuree, an agreement might usefully spell out *ex ante* the circumstances in which it would be desirable for specific assurances to be given. Alternatively, or in addition, the agreement could establish procedures by which a party that comes to suspect noncompliance can effectively communicate its concerns, prompting others to give specific assurances. From the perspective of a potential assurer, the agreement might usefully spell out agreed procedures for the transmission of information, to ensure that any assurances given will reach the appropriate officials of other governments in a timely fashion.

4. Some governments may anticipate internal opposition to the giving of assurances, fearing that the assurance strategy will be characterized as a sign of weakness by short-sighted defenders of sovereign prerogatives. In these circumstances, an explicit *ex ante* provision can be extremely useful. It allows governments to mobilize support for the assurance approach at the outset, before it is put into effect as part of a larger substantive package. With a general assurance commitment in place, governments can more easily obtain political or bureaucratic cooperation when specific assurances must be given. In addition, an *ex ante* provision may facilitate the organization of an assurance bureaucracy, providing greater efficiency and a base of support for the approach.

Hybrid verification and assurance provisions can combine these functions in a variety of ways. Assurance techniques can be tailored to feed information into the monitoring systems of other states, helping to ensure that the information transmitted will be of the appropriate scope and form and that it will be effectively communicated. States designing verification systems, on the other hand, can tailor them to the types of information other parties are willing to make available, taking advantage of the access these measures afford them and focusing their efforts on the areas where information remains inadequate. At the limit, the verifying state's desire for access can dovetail with the assuring state's desire

for effective and credible ways to convey information, leading to the creation of truly cooperative informational mechanisms in which the strategies of verification and assurance lose their separate identities.

The following table summarizes the foregoing analysis of the verification and assurance strategies and their role in international agreements.

Underlying concern	Situation where concern arises	Strategy for production of information	Provisions in international agreements
offensive defection by others	PD	verification	<ul style="list-style-type: none"> <li>- legitimate modes of observation</li> <li>- authorize 3d party monitoring</li> <li>- limit access</li> </ul>
defensive defection by others	PD Sit (only defensive incentive)	assurance	<ul style="list-style-type: none"> <li>- commit to future specific assurances</li> <li>- establish mutually acceptable forms of assurance</li> <li>- provide for timely and effective communication</li> <li>- reinforce internal arrangements</li> </ul>
		hybrid strategies	merge verification and assurance into cooperative procedures

### III. Verification and Assurance in Practice: Arms Control Agreements

This section analyzes the informational provisions in the major arms control agreements of the past 30 years. The most significant agreements include: the Limited Test Ban Treaty;<sup>95</sup> the Nuclear Non-Proliferation Treaty;<sup>96</sup> the SALT I Interim Agreement;<sup>97</sup> the Anti-Ballis-

95. Treaty Banning Nuclear Weapons Tests in the Atmosphere, in Outer Space and Underwater, Aug. 5, 1963, 14 U.S.T. 1313 [hereinafter Limited Test Ban].

96. Treaty on the Non-Proliferation of Nuclear Weapons, July 1, 1968, 21 U.S.T. 483 [hereinafter NPT].

97. Interim Agreement on the Limitation of Strategic Nuclear Arms, May 26, 1972, U.S.-U.S.S.R., 23 U.S.T. 3462 [hereinafter SALT I].

tic Missile Treaty;<sup>98</sup> the SALT II treaty;<sup>99</sup> the Intermediate-Range Nuclear Forces Agreement;<sup>100</sup> the Strategic Arms Reduction Treaty;<sup>101</sup> and the Treaty on Conventional Forces in Europe.<sup>102</sup> Several other agreements are also discussed.<sup>103</sup> A number of these agreements are multilateral, but many of the most significant were, formally or effectively, bilateral agreements between the United States and the former Soviet Union.<sup>104</sup>

---

98. Treaty for the Limitation of Anti-Ballistic Missile Systems, May 26, 1972, U.S.-U.S.S.R., 23 U.S.T. 3435 [hereinafter ABM].

99. Treaty on the Limitation of Strategic Offensive Arms, June 18, 1979, U.S.-U.S.S.R. [hereinafter SALT II], S. EXECUTIVE DOC. Y, 98th Cong., 1st Sess. 37 (1979) [hereinafter S. EXEC. DOC. Y].

100. Treaty on the Elimination of Intermediate-Range and Shorter-Range Missiles, Dec. 8, 1987, U.S.-U.S.S.R. [hereinafter INF], S. TREATY DOC. No. 11, 100th Cong., 2d Sess. 67 (1988) [hereinafter S. TREATY DOC. 11].

101. Treaty on the Reduction and Limitation of Strategic Offensive Arms, July 31, 1991, U.S.-U.S.S.R. [hereinafter START], S. TREATY DOC. No. 20, 102d Cong., 1st Sess. 1 (1991) [hereinafter S. TREATY DOC. 20].

102. Treaty on Conventional Armed Forces in Europe, Nov. 19, 1990 [hereinafter CFE Treaty], S. TREATY DOC. 8, 102d Cong., 1st Sess. 223 (1991) [hereinafter S. TREATY DOC. 8]. The parties to CFE agreed that the treaty would be applied provisionally as of July 1992 and would come into full force on ratification by the eight new states of the former Soviet Union with territory in the treaty area, as well as the original Parties. DEP'T ST. DISPATCH, July 13, 1992, at 560.

103. These include: Antarctic Treaty, Dec. 1, 1959, 12 U.S.T. 794 [hereinafter Antarctic]; Treaty on the Prohibition of the Emplacement of Nuclear Weapons and Other Weapons of Mass Destruction on the Seabed and the Ocean Floor and the Subsoil Thereof, Feb. 11, 1971, 23 U.S.T. 701 [hereinafter Seabed]; Treaty on the Prohibition of Nuclear Weapons in Latin America, Feb. 14, 1967, 22 U.S.T. 762 [hereinafter Latin America Nuclear Free Zone]; Protocol to the Treaty of July 3, 1974, on the Limitation of Underground Nuclear Weapons Tests, June 1, 1990, U.S.-U.S.S.R. [hereinafter Threshold Test Ban Verification Protocol], S. TREATY DOC. No. 19, 101st Cong., 2d Sess. 7 (1990) [hereinafter S. TREATY DOC. 19]; Protocol to the Treaty of May 28, 1976, on Underground Nuclear Explosions for Peaceful Purposes, June 1, 1990, U.S.-U.S.S.R. [hereinafter Peaceful Nuclear Explosions Verification Protocol], S. TREATY DOC. 19, at 114; Agreement on the Establishment of Nuclear Risk Reduction Centers, Sept. 15, 1987, U.S.-U.S.S.R., 27 I.L.M. 76 (1988) [hereinafter Nuclear Risk Reduction Centers].

104. Some of the agreements discussed here, such as SALT I and II, have expired. Others, however, are still in force or pending. Following the breakup of the Soviet Union, it has been important to ensure that the newly independent republics, at least four of which contain nuclear weapons, are subject to these agreements. In many cases, the doctrine of state succession in respect to treaties will resolve the issue satisfactorily. See, e.g., John B. Rhinelanders & George Bunn, *Who Is Bound by the Former Soviet Union's Arms Control Treaties?*, ARMS CONTROL TODAY, Dec. 1991, at 3. In other cases, explicit steps have been taken. For example, in May 1992, the four nuclear republics—Russia, Ukraine, Kazakhstan and Byelarus—acceded to START as independent states and agreed to adhere to the NPT as non-nuclear-weapons states; side agreements obligated the latter three republics to eliminate nuclear weapons on their territories within seven years. Protocol to the Treaty of the Reduction and Limitation of Strategic Offensive Arms, May 23, 1992, S. TREATY DOC. No. 32, 102d Cong., 2d Sess. 1, 2 (1992).

For a chronology of arms control issues involving the breakup of the Soviet Union and the formation of the Commonwealth of Independent States, see *Chronology of Commonwealth Security Issues*, ARMS CONTROL TODAY, May 1992, at 27. A full examination of these issues is beyond the scope of this article.

Together, these agreements form a body of evidence particularly suitable to the application of the theory developed in this article. For one thing, the basic assumptions of rationalist IR theory are as realistic in this area as in any other: in arms control negotiations, each government seeks assiduously to further its national interest, in this case its security, as under the self-interest assumption; and each government carefully analyzes its preferences and negotiates to advance them, as under the rationality assumption.<sup>105</sup> In addition, for much of the period under review, arms competition and arms control between the United States and the Soviet Union appeared to constitute an iterated PD interaction.<sup>106</sup> Finally, as noted above, the need for information has been a central theme of arms control policy.

### A. Verification

Offensive incentives to defect would seem to be present—and typically strong—in all competitive security relationships and certainly in those between two rival superpowers. Concern for offensive defection has in fact been prominent in East-West relations, from shortly after World War II until the present day. This concern dominated public rhetoric about arms control in the United States from the start of the Cold War through the early years of the Reagan administration.<sup>107</sup> With constantly expanding nuclear arsenals, moreover, the payoff for cooperating in the face of defection has been regarded as unacceptably low. Finally, all aspects of East-West relations, including arms control, “have labored under a burden of suspicion.”<sup>108</sup> As theory would predict in such a setting, verification devices have been prominent in arms control agreements.

#### 1. External Verification Techniques

The first modern arms control agreement, the Limited Test Ban Treaty (LTBT), did not explicitly incorporate any procedures for the production of information. Both the United States and the Soviet Union were, however, committed to the strategy of verification for monitoring compliance with the test ban, reflecting the fear of offensive defection and the lack of trust that characterized their relationship. In a classic example of verification techniques, the United States deployed around the

---

105. In addition, decisions are made at the highest levels with final input by a small number of officials, a close approximation to the assumption that states act as unitary entities. See DOWNS & ROCKE, *supra* note 8, at 92-95.

106. See DOWNS & ROCKE, *supra* note 8, at 76 (PD most characteristic arms race model, though Deadlock and Stag Hunt may also be appropriate); see also *supra* notes 23, 27.

107. Richard Perle, the chief theoretician on security policy in the Reagan administration, observed that the Soviets “care nothing for the spirit of agreements, and while it suits their purposes, little more for the letter. . . .” See James A. Schear, *Arms Control Treaty Compliance: Buildup to a Breakdown?*, INT’L SECURITY, Fall 1985, at 141, 143 (quoting Assistant Secretary of Defense Richard N. Perle, SENATE COMM. ON ARMED SERVICES, 98TH CONG., 2D SESS., SOVIET TREATY VIOLATIONS 7 (1984)).

108. See Schear, *supra* note 107, at 144.

borders of the Soviet Union both ground-based seismological monitoring stations and aircraft capable of detecting radiation from nuclear tests; it also orbited two specialized surveillance satellites in 1963.<sup>109</sup> The parties could not agree to prohibit underground testing primarily because of concerns that some tests could not be detected using current seismological techniques.<sup>110</sup>

Virtually all of the major agreements following the LTBT have explicitly authorized each party to monitor the conduct of the other parties using its own "national technical means" of verification (NTM).<sup>111</sup> This provision served to reverse the longstanding Soviet position that aerial reconnaissance over national territory, even by satellite, was a form of espionage, and to eliminate any concern that international law might prohibit satellite surveillance.<sup>112</sup> It therefore reflected the first of the three rationales for explicit verification provisions discussed above. Today, NTM remains the heart of arms control verification. The most common form of NTM is photoreconnaissance, carried out from special-

---

109. See Harold V. Argo, *Satellite Verification of Arms Control Agreements*, in *ARMS CONTROL VERIFICATION: THE TECHNOLOGIES THAT MAKE IT POSSIBLE* 290, 291 (Kosta Tsipsis et al. eds., 1986); Mark K. Lowenthal & Joel S. Wit, *The Politics of Verification*, in *VERIFICATION AND ARMS CONTROL* 153, 155-56 (William C. Potter ed., 1985).

110. It was feared that underground explosions could be muffled by conducting them in large caverns or masked by synchronizing them with earthquakes. See Kosta Tsipsis, *Introductory Note*, in *ARMS CONTROL VERIFICATION: THE TECHNOLOGIES THAT MAKE IT POSSIBLE*, *supra* note 109, at 223. Today, distant seismology techniques have been further improved. See Jack F. Evernden & Charles B. Archambeau, *Some Seismological Aspects of Monitoring a Comprehensive Test Ban Treaty*, in *ARMS CONTROL VERIFICATION: THE TECHNOLOGIES THAT MAKE IT POSSIBLE*, *supra* note 109, at 224. In a comprehensive test ban treaty, such external NTM monitoring would probably be supplemented by seismological stations located within the national territories of the parties. See Paul A. Stokes, *Unattended In-Country Stations for Seismological Verification*, in *ARMS CONTROL VERIFICATION: THE TECHNOLOGIES THAT MAKE IT POSSIBLE*, *supra* note 109, at 264.

The 1974 Threshold Test Ban Treaty, limiting underground tests to a yield of 150 kilotons, provided for monitoring by NTM, that is, by seismological stations located outside the testing state. The 1976 Peaceful Nuclear Explosions Treaty authorized some on-site monitoring. See Warren Heckrotte, *Verification of Test Ban Treaties*, in *VERIFICATION AND ARMS CONTROL*, *supra* note 109, at 63, 65. Neither agreement was ratified. In 1990, however, the United States and the Soviet Union signed verification protocols to these two agreements, and both have been submitted to the Senate for its advice and consent.

111. See ABM, *supra* note 98, art. XII, ¶ 1, 23 U.S.T. at 3443; SALT I, *supra* note 97, art. V, ¶ 1, 23 U.S.T. at 3465; SALT II, *supra* note 99, art. XV, ¶ 1, S. EXEC. DOC. Y, at 43; INF, *supra* note 100, art. XII, ¶ 1, S. TREATY DOC. 11, at 79; START, *supra* note 101, art. IX, ¶ 1, S. TREATY DOC. 20, at 33; CFE Treaty, *supra* note 102, art. XV, ¶ 1, S. TREATY DOC. 8, at 240. See also Seabed, *supra* note 103, art. III, ¶¶ 1, 5, 14 U.S.T. at 705, 706. Some agreements in areas where monitoring would be difficult do not explicitly refer to NTM, but presumably assume that NTM monitoring will take place. See, e.g., Convention on the Prohibition of the Development, Production, and Stockpiling of Bacteriological and Toxin Weapons and on Their Destruction, Apr. 10, 1972, 26 U.S.T. 583 [hereinafter Biological], art. VI, ¶ 1 (party that finds another acting in violation of treaty obligations may lodge complaint with Security Council).

112. See John L. Gaddis, *The Evolution of a Reconnaissance Satellite Regime*, in *U.S.-SOVIET SECURITY COOPERATION: ACHIEVEMENTS, FAILURES, LESSONS* 353, 358 (Alexander L. George et al. eds., 1988).

ized satellites and aircraft.<sup>113</sup> Optical photoreconnaissance is supplemented by radar and infrared imaging and other sophisticated technologies.<sup>114</sup> Other forms of NTM include the interception of electronic communications and a variety of specialized systems for monitoring activities in space and at sea.<sup>115</sup>

A few arms control agreements also reflect the second rationale for the use of explicit verification provisions by authorizing a regional or multilateral organization to perform the monitoring function. By far the most important international verification organizations (IVOs) are the International Atomic Energy Agency (IAEA)<sup>116</sup>—which monitors commitments not to use nuclear material and facilities for military purposes under bilateral agreements, both nuclear supply agreements and bilateral arms control treaties;<sup>117</sup> the multilateral Nuclear Non-Proliferation Treaty (NPT); and regional arrangements, including the Treaty of Tlatelolco, which establishes a nuclear-weapon-free zone in Latin

113. See Jeffrey Richelson, *Technical Collection and Arms Control*, in VERIFICATION AND ARMS CONTROL *supra* note 109, at 169. The capabilities of photoreconnaissance satellites are being greatly expanded by such technological advances as optics that compensate for atmospheric distortion. See generally ARMS CONTROL VERIFICATION: THE TECHNOLOGIES THAT MAKE IT POSSIBLE, *supra* note 109.

Some commercial satellite and airborne imaging systems, such as the U.S. Landsat system, could also be adapted to support verification and related activities, notably U.N. peacekeeping. See COMMERCIAL OBSERVATION SATELLITES AND INTERNATIONAL SECURITY (Michael Krepon et al. eds., 1990); ALLEN V. BANNER, OVERHEAD IMAGING FOR VERIFICATION AND PEACEKEEPING: THREE STUDIES 29-40 (1991).

114. See Eli Brookner, *Radar Imaging for Arms Control*, in ARMS CONTROL VERIFICATION: THE TECHNOLOGIES THAT MAKE IT POSSIBLE, *supra* note 109, at 135; James C. Fraser, *Infrared Surveillance and Detectors*, in ARMS CONTROL VERIFICATION: THE TECHNOLOGIES THAT MAKE IT POSSIBLE, *supra* note 109, at 166; Robert J. Bell, *Fourier Transform Spectroscopy Measurements of Infrared and Millimeter Wavelengths*, in ARMS CONTROL VERIFICATION: THE TECHNOLOGIES THAT MAKE IT POSSIBLE, *supra* note 109, at 185. Modern NTM cost "many billions of dollars per year" to operate. CONGRESSIONAL BUDGET OFFICE, 101ST CONG., 2D SESS., U.S. COSTS OF VERIFICATION AND COMPLIANCE UNDER PENDING ARMS CONTROL TREATIES 43 (1990).

115. See Richelson, *supra* note 113; Dunbar Lockwood, *Monitoring by National Technical Means*, ARMS CONTROL TODAY, Oct. 1990, at 15.

116. For useful discussions of the IAEA in the context of verification, see JAMES F. KEELEY, INTERNATIONAL ATOMIC ENERGY AGENCY SAFEGUARDS: OBSERVATIONS ON LESSONS FOR VERIFYING A CHEMICAL WEAPONS CONVENTION (1988) [hereinafter IAEA SAFEGUARDS]; James F. Keeley, *The International Atomic Energy Agency and the Non-Proliferation Treaty*, in INTERNATIONAL VERIFICATION ORGANIZATIONS (Ellis F. Morris ed., 1991) at 175 [hereinafter IVOs]. In addition to the functions described in the text, the IAEA is investigating Iraq's nuclear program under authority of the Security Council's cease-fire resolutions. See *Nuclear Proliferation: Learning from the Iraq Experience, Hearing before Senate Comm. on Foreign Relations*, 102d Cong., 1st Sess., 14 (1992) (statement of Dr. David Kay, deputy leader, International Atomic Energy Agency Action Team for Nuclear Inspections).

117. In a recent bilateral agreement, Argentina and Brazil agreed in principle in 1990 to accept full-scope safeguards on all their nuclear facilities. The two states plan to create a bilateral inspection organization, which will then negotiate a safeguards agreement with the IAEA. David Albright; *Brazil Comes in from the Cold*, ARMS CONTROL TODAY, Dec. 1990, at 13.

America<sup>118</sup>—and EURATOM.<sup>119</sup> The Treaty on Conventional Armed Forces in Europe (CFE Treaty) authorizes NATO and the former Warsaw Pact alliance, around which the agreement was structured,<sup>120</sup> to use “multinational technical means of verification.”<sup>121</sup> IVOs have also been proposed in many other contexts.<sup>122</sup>

IVOs are seen to have three principal advantages. First, it is often assumed that they will give all participating nations equal access to the data they gather. When treaties rely on NTM, in contrast, only the United States, Russia and a few other states have the technical capacity to monitor effectively on their own.<sup>123</sup> It was on this basis that France proposed a U.N. satellite monitoring agency in 1978.<sup>124</sup> Second, IVOs are often seen as more impartial, less likely to seek collateral intelligence, and less of an intrusion on national sovereignty than monitoring by other states.<sup>125</sup> Finally, in a multilateral setting, IVOs may simply be more efficient than a multiplicity of national verification systems.<sup>126</sup> The major states, however, have often opposed reliance on IVOs out of concern that they would be less demanding, flexible and competent than

118. See Latin America Nuclear Free Zone, *supra* note 103, arts. 13, 16, 22 U.S.T. at 772, 773.

119. Another independent organization, the Armaments Control Agency of the Western European Union, formerly monitored the German chemical industry for non-production of chemical weapons. For brief discussions of these agencies, see IAEA SAFEGUARDS, *supra* note 116, at 22-23; Morris, *International Verification Organizations: An Introduction*, in IVOs, *supra* note 116, at 8; F. R. CLEMINSON & E. GILMAN, A CONCEPTUAL WORKING PAPER ON ARMS CONTROL VERIFICATION 6-7, 26-27 (1986).

120. See Article-by-Article Analysis of the Treaty on Conventional Armed Forces in Europe, S. TREATY DOC. 8, *supra* note 102, at 51; CFE Treaty, *supra* note 102, arts. II, ¶ 1, IV-VII, S. TREATY DOC. 8, at 224, 229-33.

121. CFE Treaty, *supra* note 102, art. XV, ¶ 1, S. TREATY DOC. 8, at 240.

122. Most notably, the current “rolling text” of the multilateral chemical weapons convention under negotiation for many years contemplates a technical secretariat and a corps of international inspectors comparable to the IAEA. See Sutherland, *Chemical Weapons*, in IVOs, *supra* note 116, at 49; John Barrett, *Verification of a Chemical Weapons Ban: The On-Site Inspection Burden*, in ARMS CONTROL VERIFICATION AND THE NEW ROLE OF ON-SITE INSPECTION 139, 141 (Lewis A. Dunn & Amy E. Gordon eds., 1990).

123. The Open Skies agreement, described in *infra* notes 146-148, requires each state that carries out an overflight to share all the raw data collected with all other participating states. In this respect, smaller states with limited NTM capabilities are among the greatest beneficiaries of the agreement. See Jones, *infra* note 146, at 15.

124. See Lucy Stojak, *International Verification Organizations for Arms Control in Outer Space*, in IVOs, *supra* note 116, at 125, 128-29; Frank Barnaby, *Introduction*, in A HANDBOOK OF VERIFICATION PROCEDURES 1, 5 (Frank Barnaby ed., 1990).

125. See IVOs, *supra* note 116, at 5-6; IAEA SAFEGUARDS, *supra* note 116, at 21.

126. See IAEA SAFEGUARDS, *supra* note 116, at 21. To minimize the twin problems of duplication and gaps in coverage caused by independent national inspections under the CFE Treaty, NATO established a new Verification Coordination Commission with a broad mandate to coordinate national inspections, develop a common data base, redistribute unused inspection quotas, and otherwise enhance the efficiency of Western verification. See Joachim H. Schlüter, *CFE Verification and NATO*, in TOWARDS HELSINKI 1992: ARMS CONTROL IN EUROPE AND THE VERIFICATION PROCESS 21 (Heather Chestnutt and Steven Mataija eds., 1991).

national verification systems.<sup>127</sup>

## 2. *The Trend Toward More Intrusive Verification Techniques*

As states have come to demand more effective verification procedures, modern arms control agreements have built on the concept of NTM by authorizing increasingly intrusive verification techniques.<sup>128</sup> The Intermediate-Range Nuclear Forces (INF) Treaty, a product of the Reagan era, represents the watershed in this development.<sup>129</sup> The INF Treaty authorized extensive on-site inspection of the destruction of missiles, launchers and facilities, as well as challenge inspections, on short notice, of many sites designated in the agreement. The INF agreement also permits each side to conduct continuous portal-perimeter monitoring of a missile plant in which cheating was thought particularly likely. This provision calls for Soviet and American inspectors to be stationed on each other's territory until 2001.<sup>130</sup>

The Treaty on the Reduction and Limitation of Strategic Offensive Arms (START) elaborated on the pattern set in the INF agreement while applying it to a much wider range of controlled armaments, creating "the most extensive and intrusive inspection regime ever included in an arms control agreement."<sup>131</sup> Indeed, with START, on-site inspection appears to have displaced external monitoring by NTM as the primary verification mechanism. The treaty provides for 12 different types of inspection, including baseline inspections to confirm data provided on signing the agreement, inspections of conversions and eliminations of missiles and bombers, and snap "suspect-site" inspections.<sup>132</sup> In addition, START permits continuous portal-perimeter monitoring during the life of the treaty (15 years, subject to extension) at specified pro-

127. The IAEA, for example, has tended to define its mandate narrowly, and has been loath to devote greater attention to states suspected of harboring nuclear ambitions or adopting aggressive military policies than to any other party. See IAEA SAFEGUARDS, *supra* note 116, at 11, 22; Stojak, *supra* note 124, at 129-30, 134-36.

128. For a categorization of verification techniques reflecting gradations of intrusiveness, see Cleminson & Gilman, *supra* note 119, at 6-8. See generally ARMS CONTROL VERIFICATION AND THE NEW ROLE OF ON-SITE INSPECTION, *supra* note 122.

129. The Antarctic, *supra* note 103, art. VII, 12 U.S.T. at 797, Seabed, *supra* note 103, art. III, 23 U.S.T. at 704, 705, and Peaceful Nuclear Explosions Verification Protocol, *supra* note 103, authorized forms of on-site inspection, but these agreements have been much less controversial.

130. INF, *supra* note 100, art. XI, ¶ 1, S. TREATY DOC. 11, at 77 (right to conduct on-site inspections); INF, *supra* note 100, art. XI, ¶ 3, 4, 5, 7, 8; art. X, ¶ 2, S. TREATY DOC. 11, at 76-79 (types of on-site inspection); INF, *supra* note 100, art. XI, ¶ 6, S. TREATY DOC. 11, at 78 (continuous on-site monitoring of specified facilities). See Owen Greene & Patricia Lewis, *Verifying the INF Treaty and START*, in A HANDBOOK OF VERIFICATION PROCEDURES, *supra* note 124, at 215, 221, 228.

131. Letter of Submittal transmitting START from the Department of State to the President, Nov. 20, 1991, S. TREATY DOC. 20, 102d cong. 1st Sess., at vii, xiii.

132. See START, *supra* note 101, art. XI, S. TREATY DOC. 20, at 35, 36; see also START, *supra* note 101, Nov. 19, 1990, Protocol on Inspections and Continuous Monitoring Activities, S. TREATY DOC. 20, at 102.

duction facilities for mobile missiles.<sup>133</sup>

The CFE Treaty also relies on an extensive regime of on-site inspections,<sup>134</sup> as do the U.S.-Soviet agreement on the destruction of chemical weapons<sup>135</sup> and the verification protocol to the Threshold Test Ban Treaty.<sup>136</sup> IVOs have also been authorized to utilize intrusive methods of verification. For example, the IAEA utilizes in-facility containment and surveillance equipment, as well as several types of on-site inspection.<sup>137</sup> Following the discovery of Iraq's extensive nuclear weapons program, the IAEA acted to assert its formal authority, never exercised, to conduct challenge inspections of non-declared facilities.<sup>138</sup>

133. See START, *supra* note 101, art. XI, ¶ 14, S. TREATY DOC. 20, at 38; START, *supra* note 101, Nov. 19, 1990, Protocol on Inspections art. XVI, S. TREATY DOC. 20, at 184.

134. The treaty provides for four types of inspection: 1) inspections of declared sites where "objects of verification" are located, 2) challenge inspections of specified areas, 3) inspections of the "reduction" or elimination of armaments, and 4) inspections of the disarmament and certification as non-combat-capable of certain aircraft and helicopters. CFE Treaty, *supra* note 102, art. VIII, ¶ 1, XIV, S. TREATY DOC. 8, at 233-35, 240; CFE Treaty, *supra* note 102, Nov. 19, 1990, Protocol on Inspection, S. TREATY DOC. 8, at 302; CFE Treaty, *supra* note 102, Nov. 19, 1990, Protocol on Procedures Governing the Reclassification of Specific Models or Versions of Combat-Capable Trainer Aircraft into Unarmed Trainer Aircraft, S. TREATY DOC. 8, at 259; CFE Treaty, *supra* note 102, Nov. 19, 1990, Protocol on Procedures Governing the Categorisation of Combat Helicopters and the Recategorisation of Multi-Purpose Attack Helicopters, S. TREATY DOC. 8, at 277.

Interestingly, challenge inspections may be refused by the party to be inspected. If access to a specified area is denied, however, the party to be inspected "shall provide all reasonable assurance that the specified area does not contain conventional armaments and equipment limited by the Treaty." CFE Treaty, *supra* note 102, Nov. 19, 1990, Protocol on Inspection, Sec. VIII, ¶¶ 4, 9, S. TREATY DOC. 8, 322, 323-24. Assurance techniques in this instance substitute for inspection. The Protocol does not, however, specify how the necessary assurances are to be conveyed.

135. U.S.-U.S.S.R. Chemical Weapons Convention, Art. V, provides for a variety of systematic inspections and for the "continuous presence of inspectors and continuous monitoring with on-site instruments." Agreement on Destruction and Non-Production of Chemical Weapons and Measures to Facilitate the Multi-Lateral Convention on Banning Chemical Weapons, June 1, 1990, U.S.-U.S.S.R., art. V, ¶ 2, 29 I.L.M. 932, 937 (1990) [hereinafter Chemical Weapons].

136. The Verification Protocol authorizes each party to monitor tests of specified yield by (1) the hydrodynamic yield method, in which sensing instruments are lowered into a hole near the test site; (2) the seismic yield method, in which vibrations from tests are monitored at "designated seismic stations" within the territory of the testing party; and (3) on-site inspections, which largely supplement the other two. Threshold Test Ban Verification Protocol, *supra* note 103, § III, ¶ 1, §§ V-VII, S. TREATY DOC. 19, at 12, 27-74. The hydrodynamic yield method, favored by the United States, is particularly intrusive, as the testing equipment must be installed very close to the actual test site. See Wolfgang K. H. Panofsky, *Verification of the Threshold Test Ban*, ARMS CONTROL TODAY, Sept. 1990, at 3, 5. Similar arrangements are made in the verification protocol to the Peaceful Nuclear Explosions Treaty. Peaceful Nuclear Explosions Verification Protocol, *supra* note 103, §§ III, V-VII, S. TREATY DOC. 19, at 116, 124-144.

137. See Keeley, *supra* note 116, at 2, 44; Cleminson & Gilman, *supra* note 119, at 6-7.

138. The authority is contained in the safeguard agreements signed by non-nuclear-weapon state adherents to the NPT. See Jon Wolfsthal, *IAEA to Implement "Suspect Site" Inspection Powers*, ARMS CONTROL TODAY, Mar. 1992, at 27.

A related feature of INF, START, CFE and other recent agreements is their remarkable level of detail, as to both substantive obligations and provisions for verification. As Chayes and Chayes put it, the length of arms control agreements has been increasing geometrically.<sup>139</sup> The START Protocol on Inspections alone, for example, covers 181 printed pages.<sup>140</sup>

The detailed specification of substantive obligations is designed, perhaps quixotically, to make it more difficult to use ambiguities of language to rationalize defection; the approach reflects the negative assumptions behind the verification strategy.<sup>141</sup> The lengthy verification provisions stem directly from the use of complex and intrusive verification techniques. These require detailed statements of the rights of access to be accorded the personnel and equipment of the inspecting state. At the same time, parties to these innovative agreements have found it necessary to impose carefully drawn limitations on the freedom of foreign inspectors,<sup>142</sup> reflecting the third rationale for explicit verification provisions. Recent agreements have gone to compulsive extremes on this score. START, for instance, specifies the maximum number of spare flashlight bulbs and length of rulers that inspectors may carry.<sup>143</sup>

#### B. From Verification to Assurance

Perhaps surprisingly, given the conflictual nature of East-West security relations and the flood of verification rhetoric in the United States, the major arms control agreements also include a variety of rules and procedures that function as assurance devices, either alone or in combination with verification mechanisms. The following paragraphs outline these mechanisms, beginning with those best described as ancillary to verification and moving to those that more closely resemble pure assurance.

---

139. See Antonia H. Chayes & Abram Chayes, *From Law Enforcement to Dispute Settlement: A New Approach to Arms Control Verification and Compliance*, INT'L SECURITY, Spring 1990, at 147, 155.

140. See START, *supra* note 101, Nov. 19, 1990, Protocol on Inspections, S. TREATY DOC. 20, at 102.

141. See Letter of Transmittal, submitting START Treaty to Senate, Nov. 25, 1991, S. TREATY DOC. 20, at iii, iv ("[I]n order for the Treaty to work smoothly over many years, its terms must be as precise and unambiguous as possible. Neither Party should have any doubt as to the limitations and obligations that are imposed by the terms of the Treaty.") For a discussion of some of the difficulties with this approach, see GLORIA DUFFY ET AL., COMPLIANCE AND THE FUTURE OF ARMS CONTROL 2-4 (1988).

142. It has long been the Soviet position that inspection procedures should be designed to avoid the collection of "information not needed for the purpose of verification." "Verification, [the Soviets] said, is not a research program." Heckrotte, *supra* note 110, at 67. On the problem of collateral information, see James A. Schear, *Cooperative Measures of Verification: How Necessary? How Effective? in VERIFICATION AND ARMS CONTROL*, *supra* note 109, at 7, 14 [hereinafter *Cooperative Measures*].

143. START, *supra* note 101, Nov. 19, 1990, Protocol on Inspections, Annex 8, § II, ¶ B(1), S. TREATY DOC. 20, at 239, 243.

### 1. *Non-interference*

All major agreements after the LTBT explicitly provide that neither party will interfere with the other's NTM or deliberately act to conceal restricted activities from monitoring by the other side. These provisions might be called commitments to "passive acquiescence" in NTM verification.<sup>144</sup> Historically, like authorization for NTM monitoring itself, they constituted a considerable breakthrough.

The CFE negotiations and follow-on talks began to move haltingly toward an agreement permitting even low-level aerial reconnaissance flights over national territory without interference.<sup>145</sup> Then, in a remarkable manifestation of the recent political changes in Europe, 25 members of the Conference on Security and Cooperation in Europe, including Russia and other former Soviet republics, signed an "open skies" treaty in March 1992 that will allow such overflights throughout most of Europe, North America, and the Asian part of Russia.<sup>146</sup> Overflights can be made on surprisingly short notice—the proposed route need not be disclosed until 24 hours before a flight begins—and the most sensitive states, Russia and the United States, may be overflown 42 times each year. No part of the national territory may be declared off limits.<sup>147</sup> Open Skies is primarily intended as a general confidence-building measure, but may be a significant adjunct to arms control.<sup>148</sup>

---

144. See ABM, *supra* note 98, art. XII, ¶¶ 2, 3, 23 U.S.T. at 3443, 3444 (not to interfere, not to use deliberate concealment measures that impede verification); SALT I, *supra* note 97, art. V, 23 U.S.T. at 3465 (same); SALT II, *supra* note 99, S. EXEC. DOC. Y, at 43 (same); INF, *supra* note 100, art. XV, ¶ 2, S. TREATY DOC. 11, at 79 (same); START, *supra* note 100, art. IX, ¶¶ 2, 3, S. TREATY DOC. 20, at 33 (same); CFE Treaty, *supra* note 102, art. XV, ¶¶ 2, 3 S. TREATY DOC. 8, at 240, 241 (same).

145. See U.S. CONGRESS, OFFICE OF TECHNOLOGY ASSESSMENT, 102D CONG., 1ST SESS., VERIFICATION TECHNOLOGIES: COOPERATIVE AERIAL SURVEILLANCE IN INTERNATIONAL AGREEMENTS 13 (1991). The CFE Treaty provided for follow-on negotiations, dubbed CFE 1A, to focus on aerial surveillance measures as well as personnel reductions. *Id.* at 63-66; CFE Treaty, *supra* note 102, art. XIV, ¶ 6, XVIII, S. TREATY DOC. 8, at 240, 242. The aerial inspection negotiations were delayed by work on the Open Skies Agreement, discussed below, see James Macintosh, *European Arms Control Developments*, in *EAST-WEST ARMS CONTROL: CHALLENGES FOR THE WESTERN ALLIANCE* 272 (David Dewitt & Hans Rittenger eds., 1992), and were eventually replaced by that agreement. The CFE 1A Agreement, signed July 1992, is a non-legally binding limitation on military manpower in Europe. See DEP'T ST. DISPATCH, July 13, 1992, at 560-62.

146. For discussions of the Open Skies agreement, see *Gist: Open Skies Treaty*, 3 U.S. DEP'T ST. DISPATCH 257 (1992); George Leopold, *Open Skies Treaty Rings in New Era of Military Openness*, DEFENSE NEWS, April 6, 1992, at 36; Peter Jones, *Open Skies: A New Era of Transparency*, ARMS CONTROL TODAY, May 1992, at 10. For background on the negotiations, see OFFICE OF TECHNOLOGY ASSESSMENT, *supra* note 145, at 49-60; David A. Koplow, *Back to the Future and Up to the Sky: Legal Implications of "Open Skies" Inspection for Arms Control*, 79 CAL. L. REV. 421 (1991).

The Agreement allows accession by other states, and it is hoped that the eight former Soviet republics that did not sign, as well as other states in the treaty area, will accede. See Jones, *supra*, at 10.

147. See Jones, *supra* note 146, at 13. Byelarus is included with Russia for purposes of the passive overflight quota.

148. See *Gist: Open Skies Treaty*, *supra* note 146; Jones, *supra* note 146, at 13.

A commitment to passive acquiescence is clearly an adjunct to external monitoring: by itself, such a commitment conveys no information at all. Suspicious states anxious to monitor closely the actions of their treaty partners undoubtedly demand the inclusion of non-interference rules, just as they demand authorization for NTM monitoring. At the same time, however, a willingness to open oneself to external monitoring reflects a recognition that suspicious states must be given confidence if agreements are to be reached. An *ex ante* commitment to passive acquiescence anticipates the first rationale for explicit assurance provisions by giving treaty partners confidence at the outset that their verification techniques will function as effectively as possible. The assurance approach seems particularly to have characterized the Open Skies negotiations. The U.S. delegate said, for example, that the treaty "is not an information-gathering enterprise," but an instrument to ease anxieties in the region.<sup>149</sup>

## 2. Measures to Enhance Monitoring

Increasingly, the parties to arms control agreements have agreed to cooperate more actively with external monitoring, committing themselves to "cooperative measures" that go beyond mere non-interference to increase the transparency of their own activities. Accepting, or even inviting, intrusive monitoring methods like those of INF, START and CFE—which require the host state to house, transport, escort and otherwise provide for and assist a large number of inspectors—should itself be seen as a cooperative measure.<sup>150</sup> A number of agreements, however, involve even more affirmative obligations.

One rapidly evolving set of techniques involves commitments to exchange confidential data. The scope of this development is striking, given the longstanding reluctance of the United States and the Soviet Union to share military information. SALT II began by establishing an "agreed data base" of information on the number of weapons possessed by each side within the categories controlled by the agreement. In addition, each side committed itself to updating this information annually.<sup>151</sup> SALT II marked the first time the Soviets had agreed to release information on their strategic forces. The INF Treaty called for similar data exchanges, and required each party to give advance notification of many required actions, including the elimination of missiles and facili-

---

149. *Moscow Agrees to Aerial Inspection Treaty*, N.Y. TIMES, Nov. 6, 1991, at A11.

150. See *Cooperative Measures* *supra* note 142, at 7-8. The START and CFE Inspection Protocols require each party, "for the purpose of helping to ensure verification of compliance," to facilitate the work of inspectors and monitors in accordance with their terms. START, *supra* note 101, July 31, 1991, Protocol on Inspections, art. I, S. TREATY DOC. 20, at 102; CFE, *supra* note 102, Nov. 19, 1990, Protocol on Inspection, art. II, ¶ 1, S. TREATY DOC. 8, at 305.

151. See SALT II, *supra* note 99, art. XVII, ¶ 3, S. EXEC. DOC. Y, at 44; see also SALT II, *supra* note 99, art. XVI, S. EXEC. DOC. Y, at 43 (requiring advance notice of missile launches).

ties, so that monitoring efforts could be properly planned.<sup>152</sup>

START incorporates a 350-page data base with technical information about various missiles, launchers and bombers, aggregate numbers of controlled armaments, including nuclear warheads, and detailed breakdowns of arms deployments by location.<sup>153</sup> It requires the parties to update this information<sup>154</sup> and to give prior notification of over 80 different actions.<sup>155</sup> The CFE Treaty creates an extensive set of data bases, including the location and the holdings of armaments covered by the treaty for each major military unit and formation, to be periodically updated.<sup>156</sup> The treaty also establishes a complex system of prior notifications, focusing especially on the conversion and destruction of weapons.<sup>157</sup>

Data exchange and notification provisions have generally been described as adjuncts to verification,<sup>158</sup> and they do function to legitimate access to previously confidential information and improve the effectiveness of external monitoring. Yet these provisions can also be seen as assurance devices. Initial data exchanges and commitments to updating and notification signal a willingness to cooperate and thus create confidence that additional information will be forthcoming in the future. The specificity of the data provisions helps ensure that the information transmitted will be seen as acceptable. Data exchange procedures link the provision of information to foreign monitoring systems, ensuring that it will be effectively communicated. Finally, these provisions specify in advance the occasions on which further information should be supplied. In sum, the data exchange provisions of arms control agreements reflect the first three rationales for explicit assurance procedures (and probably the fourth as well).

Information transmitted by governments, of course, will often be viewed with skepticism (the fundamental dilemma of assurance). Consequently, recent arms control agreements subject data exchanges to various forms of verification. Even if the information provided is not taken

---

152. See INF, *supra* note 100, art. IX, S. TREATY DOC. 11, at 74.

153. See START, *supra* note 101, Memorandum of Understanding on the Establishment of a Data Base, S. TREATY DOC. 20, at 324.

154. START, *supra* note 101, art. VIII, ¶ 2, S. TREATY DOC. 20, at 31.

155. See Letter of Submittal for the START Treaty, *supra* note 101, S. TREATY DOC. 20, at xii. Art. VIII of the START Treaty specifies broad categories of required notifications; detailed requirements appear in relevant sections of the Treaty and its Protocols. "Notifications concerning inspections and continuous monitoring activities," Treaty art. VIII:3(i), for example, are detailed in Article III of the Protocol on Inspection. START, *supra* note 101, art. VIII, ¶ 3(i), S. TREATY DOC. 20, at 31; START, *supra* note 101, July 31, 1991, Protocol on Inspection, art. III, S. TREATY DOC. 20, at 106.

156. CFE Treaty, *supra* note 102, art. XIII, S. TREATY DOC. 8, at 239; Protocol on Existing Types of Conventional Armaments and Equipment, CFE Treaty, *supra* note 102, Nov. 19, 1990, S. TREATY DOC. 8, at 245.

157. CFE Treaty, *supra* note 102, Nov. 19, 1990, Protocol on Notification and Exchange of Information, S. TREATY DOC. 8, at 280; CFE Treaty, *supra* note 102, Nov. 19, 1990, Protocol on Inspection, Sec. X, S. TREATY DOC. 8, at 287-88.

158. See, e.g., CFE Treaty, *supra* note 102, art. XIII, S. TREATY DOC. 8, at 239.

at face value, however, data exchanges can perform valuable assurance functions, helping to alleviate suspicions that a state is trying to manipulate weapons categories or trying to cover a defection by "claiming that newly detected weapons were 'already there.'"<sup>159</sup> In addition, even questionable data exchanges may facilitate verification, especially when they disaggregate data so that states can identify non-compliance without first assembling comprehensive information.<sup>160</sup>

An especially interesting development in this area has been the conduct of experiments designed to link the verification and assurance functions of data exchanges. One important example was the 1989 U.S.-Soviet agreement to experiment with affirmative measures in connection with bilateral negotiations to limit chemical weapons (CW), an area where progress has been slowed by the difficulty of verification.<sup>161</sup>

The experiment had two phases. In the first phase, the parties were to exchange data on CW capabilities, including locations and inventories of storage and production facilities, meet to discuss the data, and carry out some on-site verification. During the second phase, they were to exchange additional, more detailed information, including plans for CW destruction, and carry out more extensive on-site inspections, including challenge inspections.

This experiment was designed in part to initiate a workable assurance mechanism by establishing in advance, through on-site inspection, that the data to be exchanged under a convention would be at least reasonably reliable and probative. Indeed that was its express purpose: the first stated goal of the agreement was "enabling each side to gain confidence in the data on CW capabilities that will be provided" by the other under a CW convention.<sup>162</sup> The parties recognized, however, that assurance devices alone were still likely to be insufficient, and they took the extraordinary step of attempting to assure each other in advance of the effectiveness of their respective verification procedures, "enabling each side to gain confidence in the inspection procedures that will be used to verify compliance" with a CW convention.<sup>163</sup>

This cooperative experiment played a significant role in the 1990 U.S.-Soviet CW agreement.<sup>164</sup> In that agreement, moreover, the parties

---

159. See *Cooperative Measures*, *supra* note 142, at 24.

160. See Letter of Transmittal from the President to the Senate, (July 9, 1991), S. TREATY DOC. 8, at IV; CONGRESSIONAL BUDGET OFFICE, U.S. CONGRESS, *THE START TREATY AND BEYOND* 104-05 (1991) [hereinafter *START AND BEYOND*].

161. Memorandum of Understanding Regarding a Bilateral Verification Experiment and Data Exchange Related to the Prohibition of Chemical Weapons, Sept. 23, 1989, U.S.-U.S.S.R., 28 I.L.M. 1438 (1989).

162. *Id.*

163. *Id.*

164. Information exchanged as part of the experiment became the baseline data for the agreement. The scheduled experimental inspections went ahead after the agreement was signed and the results influenced the detailed inspection protocol negotiated later. See S. J. Lundin & Thomas Stock, *Multilateral and Bilateral Talks on Chemical and Biological Weapons*, in *SIPRI YEARBOOK 1991: WORLD ARMAMENTS AND DISARMAMENTS* 513, 514-19.

agreed to conduct additional experiments, this time with challenge inspections of non-declared sites, to help pave the way for a multilateral CW convention.<sup>165</sup> Similar cooperative experiments were conducted to test informational arrangements for START and the Threshold Test Ban Treaty Verification Protocol, "demonstrating a level of cooperation unheard of before 'glasnost'."<sup>166</sup>

Another important approach to the problem of information reliability is the direct transmission of raw operating data, exemplified by a START provision dealing with missile flight tests.<sup>167</sup> During such tests, data is constantly transmitted to the home country by telemetry. This information is obviously reliable so long as it can be monitored directly, and such monitoring is an important facet of arms control verification. SALT I and II allowed each party to encrypt or otherwise shield its telemetry, but only to the extent that its measures did not "impede verification."<sup>168</sup> This imprecise language led to a number of disputes.<sup>169</sup> Under START, however, each party has agreed to broadcast all on-board technical measurements during each ICBM and SLBM flight test, without encryption, jamming or other interference. In addition, each party must later provide tapes of the broadcasts along with information needed to analyze them.<sup>170</sup> The State Department calls these commitments a "major breakthrough in openness and transparency."<sup>171</sup>

Recent arms control agreements provide for transparency measures going beyond the exchange of data. One important group of measures involves commitments to physically place items and activities subject to verification at agreed locations. Under START, INF, and CFE, for example, all parties have agreed to carry out the elimination or conversion of armaments at designated sites so the process can be more easily monitored, by NTM or on-site inspection.<sup>172</sup> The INF Treaty also required the parties to marshal all missiles covered by the agreement at specified deployment areas<sup>173</sup> and to move them to elimination facilities in complete organizational units, so as to enhance their visibility.<sup>174</sup>

---

165. Chemical Weapons, *supra* note 135, art. VI, ¶ 5, 29 I.L.M. at 938.

166. Panofsky, *supra* note 136, at 6.

167. START, *supra* note 101.

168. SALT I Interim Agreement, *supra* note 97, art. V; SALT II, *supra* note 99, art. XV.

169. Most importantly, the Reagan Administration charged the Soviet Union with violating SALT II by deliberately denying telemetry essential to verification. See DUFFY ET AL., *supra* note 141, at 74-88.

170. START, *supra* note 101, art. X, ¶ 4, S. TREATY DOC. 20, at 34; see also Protocol on Telemetric Information, *id.* at 312. START allows each party to conduct a limited number of flight tests annually free of the telemetry obligation. START, art. X, ¶ 6, *id.* at 34.

171. Letter of Submittal, START Message, *supra* note 155, at xii-xiii.

172. The agreements also specify the processes of conversion and elimination in considerable detail. See, e.g., START, *supra* note 101, art. VII, S. TREATY DOC. 20, at 30; START, *supra* note 101, July 31, 1991, Protocol on Procedures Governing the Conversion or Elimination of Items Subject to Treaty, S. TREATY DOC. 20, at 84.

173. INF, *supra* note 100, art. VIII, S. TREATY DOC. 11, at 73-74.

174. INF, *supra* note 100, art. X, ¶ 3, S. TREATY DOC. 11, at 76.

START utilizes other locational devices, especially in connection with its limits on mobile missiles, which are inherently difficult to monitor. The parties to the treaty agreed to base all mobile missile launchers within designated "restricted areas." These areas must be in designated regions, may be no larger than five square kilometers, and may not contain structures capable of housing more than the approved number of missiles. Missiles and launchers may be removed from restricted areas only for stated purposes, with limits on the frequency and duration of such removals.<sup>175</sup> Non-deployed missiles and launchers of all kinds must be located only at specified sites, and the sites themselves must be located at least 100 kilometers from the nearest base or deployment area.<sup>176</sup> The CFE Treaty, which also limits relatively small and mobile items like tanks and aircraft, utilizes similar devices. Restricted items not held by active military units, for example, must be kept at "designated permanent storage sites," with continuous perimeter fences and a limited number of gates.<sup>177</sup>

INF and START incorporate even more striking cooperative measures. In order to "enhance observation by national technical means," the parties to the INF Treaty agreed that, until 1991, each side could request the other, six times each year, to open the roofs of the missile garages at a fully operational missile base where restricted missiles could be easily concealed, remove all missiles and launchers from the garages, and display them in the open for a defined period of time.<sup>178</sup> Similarly, under START each party may request the other, seven times per year, to display in the open for a defined period, without any concealment measures, the mobile missile launchers located at a designated basing area or the heavy bombers at a designated air base. In the case of road-mobile launchers, the roofs of the storage structures must be opened throughout the display.<sup>179</sup> Much the same approach is used to facilitate monitoring of certain nuclear missile submarines converted for special purposes.<sup>180</sup>

---

175. START, *supra* note 101, art. VI, ¶¶ 9, 11, S. TREATY DOC. 20, at 27. The restrictions stated apply to "road-mobile" missile launchers; parallel restrictions apply to "rail-mobile" launchers. START, art. VI, ¶ 4, *id.* at 28.

176. See START, *supra* note 101, art. IV, ¶ 9, 11, *id.* at 17-20; Letter of Submittal, *id.* at x. The treaty also limits the number of ballistic missile submarines in drydock within 5 km. of a submarine base. START, art. IV, ¶ 6, *id.* at 16. For discussion of measures of this type, see *Cooperative Measures*, *supra* note 142, at 22-23.

177. CFE Treaty, *supra* note 102, art. II, ¶ 1(H), S. TREATY DOC. 8, at 226; CFE Treaty, *supra* note 102, art. IVm ¶¶ 1-2, S. TREATY DOC. 8, at 229; CFE Treaty, *supra* note 102, art. X, S. TREATY DOC. 8, at 236-38; CFE, *supra* note 102, art. XI, S. TREATY DOC. 8, at 238-39.

178. INF, *supra* note 100, art. XII, ¶ 3, S. TREATY DOC. 11, at 79-80. This procedure involved some real risk, not only the largely theoretical one of exposing missiles to attack, but the more substantial risk of conveying information about non-regulated weapons systems.

179. START, *supra* note 101, art. XII, ¶ 1(a), S. TREATY DOC. 20, at 38.

180. Thirty-third Agreed Statement, START, Annex on Agreed Statements, *supra* note 101 at 48, 61.

Treaty provisions like these are often touted as part of a strict verification strategy. Yet they also represent striking commitments to provide concrete future assurances in a direct and unmistakable fashion, linked to the specific needs and verification mechanisms of treaty partners. Even more clearly than data exchanges, these transparency commitments can be seen as reflecting the principal rationales for explicit assurance provisions.

### 3. *Active Cooperation*

During the SALT process, the United States and the Soviet Union began to develop even more active cooperative measures designed to demonstrate compliance. SALT II, for example, set numerical limits on heavy bombers capable of carrying certain missiles, but permitted the parties to meet those limits by converting bombers to other uses. This approach created a significant verification problem: bomber-style aircraft are extremely versatile, and a bomber differs from a tanker only in its ability to carry weapons.<sup>181</sup> In order to reach agreement on the limitation of heavy bombers, then, it was necessary to devise a way for each side to show the other that its converted bombers had in fact been properly modified. The agreed solution was to incorporate "functionally related observable differences" (FRODS), physical features that would demonstrate to an external observer using NTM that an aircraft was incapable of carrying restricted weapons.<sup>182</sup> In essence, each party undertook to design and construct its bombers in ways that would provide assurances to the other. The Carter administration considered a similar approach for the design of MX missile bases, as a way to assure the Soviets that its missiles were not being moved in a destabilizing way.<sup>183</sup>

START also requires that various types of heavy bombers, including heavy bombers equipped for various nuclear and non-nuclear weapons, converted heavy bombers, and trainers—as well as other new and converted armaments—be made "distinguishable," essentially through the incorporation of FRODS, and provides for exchanges of information

---

181. See Dean A. Wilkening, *Monitoring Bombers and Cruise Missiles*, in VERIFICATION AND ARMS CONTROL, *supra* note 109, at 107, 109-10.

182. See *id.* at 110-11; SALT II, *supra* note 99, June 18, 1979, Agreed Statements Regarding the Treaty, S. EXEC. DOC. Y, at 49. Unfortunately, because of the very structural characteristics that make them necessary, relatively few FRODS are probative, easy to observe by NTM, and difficult to reverse. See Wilkening, *supra* note 181, at 110.

183. SALT II restricted "mobile" missiles as an aid to NTM monitoring. The U.S., however, wished to attain some of the stability associated with mobility, while assuring the Soviet Union that the informational advantages of static siting had not been lost. The United States planned to shuttle each missile among several silos, but to enclose each set of silos with high earthworks and to destroy the access roads once the missiles were inside the enclosure. In addition, all the silos in each set would periodically be opened to demonstrate that there was only a single missile inside. See *Cooperative Measures*, *supra* note 142, at 18-19.

on weapon types and configurations.<sup>184</sup> In line with its general emphasis on on-site verification, however, the treaty also establishes several inspection regimes designed to confirm identification and distinguishability.<sup>185</sup> The CFE Treaty also relies on inspections to confirm the conversion of combat aircraft and helicopters.<sup>186</sup>

Under the START procedures for the reduction of mobile missiles, the parties agree to give each missile a unique identification number, and to provide the other with information on the assignment of these numbers. This system, together with the agreed data base and on-site inspection of elimination procedures, should significantly facilitate verification of compliance with the treaty's numerical limits on mobile ICBMs.<sup>187</sup> A similar serial number procedure is utilized under the CFE Treaty.<sup>188</sup> Future arms control agreements may rely on identification measures of this kind, incorporating advanced technologies, more than on FRODS. Examples include the installation of active or passive radiation-detection devices on armaments that states have agreed to keep non-nuclear, and the affixing of unique, tamper-proof tags to missiles, warheads and other armaments subject to agreed numerical restrictions.<sup>189</sup>

#### 4. Confidence-Building Measures

The major arms control agreements are embedded in a growing network of supporting agreements and institutions designed to avoid accidental breakdowns due to misinformation or the misinterpretation of ambiguous data, as well as to minimize events that could place unnecessary stress on the arms control consensus.<sup>190</sup> These mechanisms, designed to prevent premature or erroneous defensive defections, rep-

184. START, *supra* note 101, art. III, ¶ 9(e), S. TREATY DOC. 20, at 11; *see also* Annex on Terms and Their Definitions, *id.* at 67, 70 (definition of "distinguishable"). In a separate letter agreement, the United States also agreed to incorporate specified physical changes in B-1 bombers equipped with long-range nuclear cruise missiles. Exchange of Letters between U.S. Ambassador Linton F. Brooks and U.S.S.R. Ambassador Yuri K. Nazarkin, July 31, 1991, *id.* at 706-07.

185. START, *supra* note 101, art. XI, ¶¶ 11-13, *id.* at 36-38.

186. CFE Treaty, *supra* note 102, Nov. 19, 1990, Protocol on Procedures Governing the Reclassification of Specific Models or Versions of Combat-Capable Trainer Aircraft into Unarmed Trainer Aircraft, S. TREATY DOC. 8, at 259; CFE Treaty, *supra* note 102, Nov. 19, 1990, Protocol on Procedures Governing the Categorisation of Combat Helicopters and the Recategorisation of Multipurpose Attack Helicopters, S. TREATY DOC. 8, at 277.

187. START, *supra* note 101, art. IX, ¶ 4, S. TREATY DOC. 20, at 33; *see also* START, *supra* note 101, July 31, 1991, Protocol on Inspections and Continuous Monitoring Activities Relating to Treaty, S. TREATY DOC. 20, at 102.

188. *See* CFE Treaty, *supra* note 102, Nov. 19, 1990, Protocol on Inspection, § X(7)-(8), S. TREATY DOC. 8, at 326.

189. *See* Steven Fetter & Thomas Garwin, *Using Tags to Monitor Numerical Limits in Arms Control Agreements*, in TECHNOLOGY AND THE LIMITATION OF INTERNATIONAL CONFLICT 33 (Barry M. Blechman ed., 1989); START AND BEYOND, *supra* note 160, at 106-10.

190. *See* Barry M. Blechman, *Efforts to Reduce the Risk of Accidental or Inadvertent War*, in U.S.—SOVIET SECURITY COOPERATION: ACHIEVEMENTS, FAILURES, LESSONS 466

resent important manifestations of the assurance strategy. Unlike the transparency measures previously discussed, moreover, these confidence-building measures are not tied to the verification procedures of other states.

One set of devices is designed to improve official channels for the transmission of information, with the special aim of ensuring that explanations and assurances can be quickly and surely conveyed when ambiguous events occur; these provisions reflect the third rationale for explicit assurance provisions. The most familiar is the so-called Hot Line, established in 1963 and twice modernized.<sup>191</sup> Nuclear Risk Reduction Centers were established shortly before the signing of the INF treaty in 1987 to transmit notifications and information.<sup>192</sup> INF and START, as well as other recent agreements, rely on these centers for continuous and secure communications.<sup>193</sup>

A second set of devices seeks directly to avoid the accidental outbreak of war. The most notable of these is the 1971 Accidents Agreement.<sup>194</sup> Under this agreement, the parties agree to notify each other immediately (1) in case of an accidental or otherwise unexplained nuclear detonation; (2) in case of the detection by missile warning systems of an unidentified object, if those incidents could create a risk of war; and (3) in advance of missile launches expected to leave national territory in the direction of other party. More generally, in connection with all "unexplained nuclear incidents," the parties agree to act to reduce the possibility that their actions will be misinterpreted, especially by providing information to each other.

A third set of devices, negotiated at the Conference on Security and Cooperation in Europe, was designed to build confidence within the heavily armed European area in the context of conventional as well as nuclear forces.<sup>195</sup> The 1975 Helsinki Final Act contained limited "confidence and security-building measures" (CSBMs) of this sort, including

---

(Alexander L. George et al. eds., 1987) [hereinafter *SECURITY COOPERATION*]; *CONFIDENCE BUILDING IN EAST-WEST RELATIONS* (Karl E. Birnbaum ed., 1982).

191. See Memorandum of Understanding Between USA and USSR Regarding the Establishment of a Direct Communications Link, June 20, 1963, U.S.-U.S.S.R., 14 U.S.T. 825; Agreement Between the United States of America and the Union of Soviet Socialist Republics on Measures to Improve the USA-USSR Direct Communications Link, Sept. 30, 1971, U.S.-U.S.S.R., 10 I.L.M. 1174; USSR-USA: Memorandum of Understanding on the Direct Communications Link, July 17, 1984, 23 I.L.M. 1393.

192. Nuclear Risk Reduction Centers, *supra* note 103.

193. INF, *supra* note 100, art. IX, ¶ 2, S. TREATY DOC. 11, at 74; START, *supra* note 101, art. VIII, ¶ 2, S. TREATY DOC. 20, at 31; Threshold Test Ban Verification Protocol, *supra* note 103, S. TREATY DOC. 19, at 11. For background on the Centers, see Stephen I. Griffiths, *The Implementation of the INF Treaty*, in SIPRI YEARBOOK 1990: WORLD ARMAMENTS AND DISARMAMENTS 443, 453.

194. Agreement on Measures to Reduce the Risk of Outbreak of Nuclear War, Sept. 30, 1971, U.S.-U.S.S.R., 22 U.S.T. 1950.

195. Many of these agreements, of course, were entered into well before the recent CFE Treaty, the related withdrawals of forces, and the dissolution of the Warsaw Pact and the Soviet Union. New CSBMs for Europe are still being considered, however.

advance notice of very large military maneuvers.<sup>196</sup> The 1986 Stockholm review conference greatly expanded these measures, including pre-notification of smaller and more varied maneuvers, longer notice periods—up to 2 years for very large exercises—and mandatory observers. The Stockholm Document also authorized challenge inspections in case of suspected violations.<sup>197</sup>

In 1989, the United States and USSR agreed to limit certain dangerous military activities, such as the use of lasers, in close proximity to each other's armed forces, national borders, and designated Special Caution Areas. The agreement also provides for secure communications in the event such incidents occur, and commits the parties to provide appropriate clarifying information.<sup>198</sup> Shortly thereafter, they agreed to expand pre-notification requirements to major exercises involving heavy bombers.<sup>199</sup>

Most recently, in 1990, 34 states of the CSCE adopted the Vienna Document, reiterating and expanding upon the Stockholm CSBMs.<sup>200</sup> The Vienna Document calls for expanded information exchanges, including plans for new weapons systems and military budgets; provides for consultation and assurance procedures in cases of unusual military activity and "hazardous incidents;" expands the right of inspection; establishes a new communications network; and reduces the thresholds for pre-notification.<sup>201</sup> These activities are supported by the Conflict Prevention Centre of the CSCE, established under the Charter of Paris for a New Europe.<sup>202</sup>

---

196. Conference on Security and Cooperation in Europe: Final Act, Aug. 1, 1975, 14 I.L.M. 1292, 1297 (incorporating Document on confidence-building measures and certain aspects of security and disarmament).

197. Document of the Stockholm Conference on Confidence- and Security-Building Measures and Disarmament in Europe, Sept. 19, 1986, reprinted in SIPRI YEARBOOK 1987: WORLD ARMAMENTS AND DISARMAMENTS 355 (1987) (entered into force Jan. 1, 1987). For discussions, see CARL C. KREHBIEL, CONFIDENCE- AND SECURITY-BUILDING MEASURES IN EUROPE: THE STOCKHOLM CONFERENCE (1989); James E. Goodby, *The Stockholm Conference: Negotiating a Cooperative Security System for Europe*, in SECURITY COOPERATION, *supra* note 190, at 144; John Borawski et al., *The Stockholm Agreement of September 1986*, 30 ORBIS 643 (1987). For a discussion of the relationship between the CSBM and conventional forces negotiations, see James Macintosh, *Confidence-Building Processes — CSCE and MBFR: A Review and Assessment*, in EAST-WEST ARMS CONTROL: CHALLENGES FOR THE WESTERN ALLIANCE 119 (David Dewitt & Hans Rattinger eds., 1992).

198. Agreement Between Government of USA and Government of USSR on Prevention of Dangerous Military Activities, June 12, 1989, U.S.-U.S.S.R., 28 I.L.M. 877 (entered into force Jan. 1, 1990).

199. Agreement on Reciprocal Advance Notification of Major Strategic Exercises, Sept. 23, 1989, U.S.-U.S.S.R., 28 I.L.M. 1436.

200. Vienna Document 1990 of the Negotiations on CSBMs Convened in Accordance with the Relevant Provisions of the Concluding Document of the Vienna Meeting of the CSCE, Nov. 17, 1990, reprinted in SIPRI YEARBOOK 1991: WORLD ARMAMENTS AND DISARMAMENTS 475 (1991).

201. *Id.* at 477; Lee Feinstein, *CSCE Structure Set*, ARMS CONTROL TODAY, Dec. 1990, at 24.

202. Charter of Paris for a New Europe, Nov. 21, 1990, 30 I.L.M. 190, 207 [hereinafter Charter of Paris]; Supplementary Document to Give Effect to Certain Provisions

Proponents of CSBMs sometimes fret that the agreements reached have been limited and incomplete, but by now there are numerous and increasingly elaborate confidence-building arrangements. Almost all are designed so that Party A can assure Party B that what it observes is not a threat, or so that Party A structures its own activities to prevent Party B from perceiving any threat, often both. The parties to these CSBMs have recognized the shared risk of unintentional conflict based on defensive defection, and have acted to create a climate of greater confidence. In tone and structure, many of them are close to pure assurance devices.<sup>203</sup> Unfortunately, like all such measures, CSBMs can be used to mislead.<sup>204</sup> They may also lull states into unwarranted complacency, substituting false negatives for false positives.<sup>205</sup> As a result, CSBMs like those in the Stockholm and Vienna Documents often incorporate some verification procedures, creating a hybrid informational mechanism with assurance as the dominant approach.

### 5. *The SCC*

The SALT I and II agreements, including the ABM treaty, created an institution called the Standing Consultative Commission (SCC).<sup>206</sup> The SCC is primarily a forum for consultation between the parties, not an independent institution.<sup>207</sup> It has various administrative functions, such

---

Contained in the Charter of Paris for a New Europe, 30 I.L.M. at 209, 212. See Adam D. Rotfeld, *New Security Structures in Europe: Concepts, Proposals, and Decisions*, in SIPRI YEARBOOK 1991: WORLD ARMAMENTS AND DISARMAMENTS 585, 586 (1991).

203. Particular CSBMs, however, may increase the security of certain states more than that of others. The NATO states, for example, argue that they need to conduct a significant number of military exercises, especially reinforcement exercises, because of the distance of the United States from Europe. NATO states accordingly oppose limits on exercises per se during CSBM negotiations, favoring the kind of transparency measures adopted at Helsinki, Stockholm and Vienna. See Robert D. Blackwill & Jeffrey W. Legro, *Constraining Ground Force Exercises of NATO and the Warsaw Pact*, 14 INT'L SECURITY, Winter 1989-90, at 68.

204. A Rand Corporation study of NATO proposals for the Stockholm conference, for example, expressed concern that dedicated communication links could be used to transmit disinformation and that pre-notified exercises could be used as cover for offensive military actions. See Y. BEN-HORIN ET AL., BUILDING CONFIDENCE AND SECURITY IN EUROPE 22-25, 30-31 (1986).

CSBMs are also criticized as ineffective. A later Rand study, based on a crisis simulation, found that the Stockholm CSBMs did not seriously limit either side, even when they moved military forces as part of the evolving crisis; and did little to promote confidence or help the two sides clarify their intentions. In particular, forecasts of military exercises were manipulated to create surprise, minimum notification requirements became ceilings on communication, and communication links were used to convey threats and attempt intimidation. See JAMES P. KAHAN ET AL., TESTING THE EFFECTS OF CONFIDENCE- AND SECURITY-BUILDING MEASURES IN A CRISIS: TWO POLITICAL-MILITARY GAMES (1987).

205. BEN-HORIN, *supra* note 204, at 11-12, 18-21, 22-25.

206. ABM, *supra* note 98, art. XIII, 23 U.S.T. 3444; SALT I, *supra* note 97, art. VI, 23 U.S.T. 3466; SALT II, *supra* note 99, art. XVII, S. EXEC. DOC. Y, at 44.

207. For discussions of the SCC, see Robert W. Buchheim & Phillip J. Farley, *The U.S.-Soviet Standing Consultative Commission*, in SECURITY COOPERATION, *supra* note 190, at 254; DUFFY ET AL., *supra* note 141, at 163-84; Antonia H. Chayes & Abnan Chayes,

as elaborating detailed procedures and considering measures to increase treaty viability. These functions reflect a collaborative, managerial approach to coexistence under complex long-term agreements, with their inevitable ambiguities, hidden weaknesses, and changes in context.<sup>208</sup>

The most remarkable thing about the agreements creating the SCC, though, is their approach to dispute resolution. In Article XIII:1(a) of the ABM treaty, first of all, the parties agree, within the framework of the SCC, to "consider questions concerning compliance . . . and related situations which may be considered ambiguous."<sup>209</sup> This paragraph resembles a standard consultation provision, which binds each party to consider complaints about compliance and similar matters brought to its attention by others. Even a provision of that kind provides an agreed procedure for alerting states to the concerns of others, so that they may consider providing assurances. The SALT provisions, though, have an unusually collaborative tone: the parties agree to consider issues jointly, not adversarially, and to consider ambiguous situations, not only complaints.

More notable still is paragraph 1(b), under which the parties agree to "provide on a voluntary basis such information as either . . . considers necessary to assure confidence in compliance" with the treaty.<sup>210</sup> This language makes clear that the SCC was intended to be a channel for the provision of information—direct, expeditious and confidential—to forestall defensive defections or losses of confidence caused by failures in NTM monitoring, ambiguities in observed data, incomplete evidence, and the like. It was designed, in short, as an instrument for the provision of assurances, even when others have not requested them.

By most accounts, moreover, the SCC actually performed this function, at least to some extent, during the first years of SALT, roughly 1972-80. The most positive assessment is contained in a report submitted to Congress by the Carter Administration State Department in 1979, in anticipation of debate on SALT II.<sup>211</sup> It discusses the disposition of issues raised in the SCC by both sides. Neither side, it appears, went so far as to initiate clarification of ambiguities in its own conduct; each responded only to issues raised by the other. The Soviet Union did,

---

*Living Under a Treaty Regime: Compliance, Interpretation, Adaptation*, in ANTONIA H. CHAYES & E. PAUL DOTY, *DEFENDING DETERRENCE* 197 (1989).

208. For a broader discussion of the managerial mode, see Chayes & Chayes, *supra* note 139.

209. ABM, *supra* note 98, art. XIII, ¶ 1(a), 23 U.S.T. 3444.

210. *Id.* at art. XIII, ¶ 1(b), 23 U.S.T. 3444.

211. See *Compliance with SALT I Agreements*, U.S. DEP'T OF ST., Special Rep. No. 55, July 1979.

While sharing a generally positive assessment, Buchheim and Farley note that some issues discussed in the report were actually resolved in the SALT II negotiations, not the SCC; that one issue was resolved so slowly as to suggest intentional delay by the Soviets; and that Reagan administration officials later used some of the same issues as examples of Soviet non-compliance. See Buchheim & Farley, *supra* note 207, at 260-61.

however, acknowledge voluntarily that it would not meet a certain deadline, and negotiated conditions for an extension.

In response to issues raised by the United States, the Soviets often responded with considerable explanatory data; in other cases they ceased the questioned activity. For its part, the United States apparently went to substantial lengths to clarify ambiguous situations questioned by the Soviets, even though some in the United States viewed the Soviet questions as mere "counterpunching" designed to maintain an appearance of symmetry.

For example, the United States retained a number of old Titan and Atlas missiles. The Soviets inquired whether these missiles had been sufficiently dismantled so that they could not be quickly reactivated, as SALT I required. The United States responded that the missiles were not subject to the agreement, as they had been deactivated before it was signed, and produced evidence of the date of deactivation. The United States went further, however, in the interest of assuring the Soviet Union: it provided information on the condition of the missiles demonstrating that they could not be reactivated.<sup>212</sup>

In another example the Soviets expressed concern that new U.S. Pave Paws early warning radars on the East and West coasts could become part of an unlawful radar base for ABM defense. The United States responded that the radars were for early warning only. Again, however, the United States went further, providing technical information to demonstrate that the Pave Paws radars were not suitable for ABM purposes.<sup>213</sup>

The clearest failure of the SCC was the well-known case of the Krasnoyarsk radar, a situation closely parallel to that of the Pave Paws radars. By all accounts, even after the United States noted that the situs of the Krasnoyarsk radar, a considerable distance from the "periphery" of Soviet territory, appeared to violate the ABM treaty, the Soviets failed to come forward with clarifying information or even a persuasive rationale.<sup>214</sup> This failure significantly weakened the U.S. commitment to SALT and the SCC.<sup>215</sup> Much later, the Soviets invited American legislators, expert staff and journalists to tour the Krasnoyarsk site, at the risk of conveying useful military information.<sup>216</sup> In fact, this visit revealed the radar's poor quality, convincing many in the United States that Krasnoyarsk was not a significant treaty violation. Underscoring the Soviets'

---

212. See *Compliance with SALT I Agreements*, *supra* note 211.

213. *Id.*

214. For criticism of the rationale actually put forward—that the radar was intended for space tracking—see Schear, *supra* note 107, at 154-59.

215. For a discussion of this and other compliance disputes under arms control treaties, see DUFFY ET AL., *supra* note 141, at 105-12, 147-59; Gloria Duffy, *Conditions That Affect Arms Control Compliance*, in SECURITY COOPERATION, *supra* note 190, at 270; Gary L. Guertner, *Three Images of Soviet Arms Control Compliance*, 103 POL. SCI. Q. 321 (1988).

216. See William J. Broad, *Soviet Radar on Display: The Benefits to Moscow of Opening the Site Outweigh any Political and Military Losses*, N.Y. TIMES, Sept. 9, 1987, at A1.

failure to utilize the SCC as an assurance device, Soviet Foreign Minister Eduard Shavardnadze later admitted that Krasnoyarsk was a violation and the radar was dismantled.<sup>217</sup>

Based partly on episodes like Krasnoyarsk and partly on more fundamental attitudes, the Reagan administration initially adopted a more suspicious and aggressive attitude toward Soviet compliance, criticizing Soviet actions publicly instead of raising issues in the SCC.<sup>218</sup> The Department of Defense vividly urged the abandonment of the SCC as an "Orwellian memory hole" into which issues of concern to the United States had been "dumped like yesterday's trash."<sup>219</sup> The INF treaty did not utilize the SCC, but established a new entity for consultations, the Standing Verification Committee (SVC).<sup>220</sup> The SVC retained most of the administrative functions of the SCC. In terms of dispute resolution, however, the change of name marked a change to a more adversarial approach.<sup>221</sup> The stated function of the SVC was "to resolve questions related to compliance;" references to joint consideration of ambiguous situations and to providing information to create confidence were removed, substantially eliminating the assurance function.

START created yet another advisory body called the Joint Compliance and Inspection Commission (JCIC).<sup>222</sup> Its approach and procedures are modeled on those of the SVC, but it added time limits, a provision for special sessions to deal with urgent concerns "relating to compliance," and other refinements designed to tighten up the procedures.<sup>223</sup> The "urgent concern" provision, however, outlines a special form of on-site inspection, called the "visit with special right of access," which the parties may agree to utilize on a case-by-case basis as a way of resolving disputes over ambiguous situations.<sup>224</sup> This hybrid of verification and assurance would be an ideal way of dealing with issues like the Krasnoyarsk radar.

217. See *Soviets Begin Dismantling Krasnoyarsk*, ARMS CONTROL TODAY, July-Aug. 1990, at 27.

218. See, e.g., DUFFY ET AL., *supra* note 141, at 163-84; Chayes & Chayes, *supra* note 139, at 160-61; Kenneth L. Adelman, *Why Verification is More Difficult (and Less Important)*, 14 INT'L SECURITY, Spring 1990, at 141; Guertner, *supra* note 215. The same change in attitude led to the intense focus on on-site inspection in the INF Treaty and subsequent agreements.

219. DUFFY ET AL., *supra* note 141, at 31 (quoting *Responding to Soviet Violations Policy Study*, Memorandum to the President, Office of the Secretary of Defense, Nov. 13, 1985, at 9).

220. INF, *supra* note 100, art. XIII, ¶ 1, S. TREATY DOC. 11, at 80.

221. See Chayes & Chayes, *supra* note 139.

222. START, *supra* note 101, art. XV, S. TREATY DOC. 20, at 45-46; START, *supra* note 101, July 31, 1991; Protocol on the Joint Compliance and Inspection Commission Relating to Treaty, S. TREATY DOC. 20, at 319 [hereinafter Protocol of JCIC].

223. START, *supra* note 101, art. XV (a), *id.* at 45; Protocol on JCIC, *supra* note 222, art. II, ¶ 1; *id.*, art. III, ¶ 1. See Article-by-Article Analysis of Treaty Text, *id.*, at 735, 1049-54 [hereinafter Treaty Analysis].

224. Treaty Analysis, *id.* at 1051-52. The specific procedures for such visits are left for ad hoc agreement, so that the need for information and the need for confidentiality can be appropriately balanced.

The CFE Treaty, with its twenty-two parties, created a separate institution, the Joint Consultative Group (JCG).<sup>225</sup> The JCG was given an expanded set of collaborative administrative tasks.<sup>226</sup> In addition, perhaps because of the treaty's multilateral character, the dispute settlement mandate of the JCG harkens back somewhat to that of the SCC. Although CFE does not contain the extraordinary commitment to assurance found in SALT I and II, the parties do agree that they will "seek to resolve ambiguities and differences of interpretation" within the JCG, as well as "addressing" questions relating to compliance and considering other matters of dispute.<sup>227</sup> Unfortunately, the early experience of the JCG was not encouraging: the military and other conservative factions in the Soviet Union, highly influential in late 1990 and early 1991, again failed to take advantage of the opportunity to assure the West in connection with a number of important ambiguities.<sup>228</sup> It remains to be seen whether, in the post-Soviet era, the parties to CFE will take advantage of the assurance opportunities provided in the treaty.

## Conclusion

This Article demonstrates that the concepts of verification and assurance, and the rational design approach from which they are derived, can provide rich insights into the strategic relationships among states and the functioning of international agreements.

On the first point, it seems clear that over the last 30 years the United States, the Soviet Union and its successors, and the other parties to the major arms control agreements have seen offensive defection by treaty partners as a significant threat, albeit one that waxed and waned over time. In addition, those states, not surprisingly, have seen their own payoffs for cooperation in the face of defection as unacceptably low. These strategic concerns have been heightened by serious information problems stemming from the general uncertainty of international relations, the special secrecy in which military activities are conducted, and the atmosphere of suspicion that characterized the East-West rivalry.

---

225. CFE Treaty, *supra* note 102, art. XVI, S. TREATY DOC. 8, at 241-42; CFE Treaty, *supra* note 102, Nov. 19, 1990, Protocol on the Joint Consultative Group, S. TREATY DOC. 8, at 330.

226. These include resolving technical matters arising during implementation in such a way as to work toward common technical practices and working out measures to safeguard sensitive collateral information during data exchanges and inspections. CFE Treaty, *supra* note 102, art. XVI, ¶ 2, S. TREATY DOC. 8, at 281.

227. *Id.*

228. Baseline inspections under CFE demonstrated significant disparities in the information the Soviets had provided upon signing the treaty. Although the treaty allows 90 days for clarification of such issues, regular meetings of the JCG were unable to resolve them within that time. In addition, the Soviets moved a large number of armaments outside the treaty area shortly before the agreement was signed; although legal in themselves, these moves created new ambiguities that were not easily resolved. See Jane M. Sharp, *Conventional Arms Control in Europe*, in SIRPI YEARBOOK 1991: WORLD ARMAMENTS AND DISARMAMENT 407, 428-33 (1991).

As a result, the demand for effective verification measures has been a pervasive theme in arms control negotiations throughout this period. Verification devices grew more extensive, intrusive and costly as the years progressed; these developments recently accelerated as the substantive agreements became more sweeping. Verification concerns also became increasingly integral to the substantive obligations of arms control treaties: under the most recent agreements, the entire life cycle of restricted armaments—their design, testing, deployment, movement, and elimination—and other significant military activities are planned with verification in mind.

What is less well known, however, is the concomitant growth of assurance devices as an integral part of arms control agreements. The on-site inspection schemes normally advertised as tough verification measures, for example, would not be possible without: (1) the agreement of other parties to open their borders and military sites to foreign inspectors and to support their work; (2) elaborate data exchanges and systems of advance notification; and even, in some cases, (3) the substantive modifications of military activities to accommodate foreign observation techniques.<sup>229</sup> Similarly, although external monitoring technologies have grown incredibly powerful, NTM verification could not achieve its full potential if states were not willing to locate military equipment and conduct sensitive activities at agreed sites, to open the roofs of missile silos, to broadcast missile telemetry without encryption, and to design bombers with distinct identifying characteristics.

In addition to these measures, designed specifically to increase transparency, communicate information and enhance verification, confidence in continued cooperation has slowly been increased by agreements to limit threatening activities, to transmit prompt explanations of ambiguous situations, to give advance notice of major military activities, to permit unrestricted aerial surveillance, and to accept observers at military exercises. The degree of openness which measures like these have created, especially in Europe, would simply have been unthinkable at the time of the LTBT. In sum, the East-West security relationship has not been wholly adversarial; each side has acted, to a greater or lesser degree at different times and on different issues, to reduce the other's incentives for defensive defection and maintain a cooperative relationship.

What of the future? It is possible to make a few brief predictions and suggestions. First, the INF model—verification through NTM and on-site inspections, supported by data exchanges, pre-notification obligations, locational restrictions, active cooperation measures and other assurance devices—has been highly refined and well established through INF, START and the CFE treaty. The same model, with appropriate

---

229. For example, underground nuclear test procedures must be specially modified to make the hydrodynamic yield method of on-site measurement, utilized under the verification protocol to the Threshold Test Ban Treaty, effective. See Panofsky, *supra* note 136.

modifications, is likely to be followed in the "START II" treaty between the United States and Russia<sup>230</sup> and, for the near future, in other agreements between the Western allies and the nations of the former Soviet Union.

If a climate of greater trust between East and West were to develop, however, assurance might come to play an even more significant role.<sup>231</sup> In September and October 1991, for example, the United States and the Soviet Union reached independent, though coordinated, decisions to eliminate large classes of tactical nuclear weapons.<sup>232</sup> Without some degree of trust, or some confidence that the other side has cooperative preferences on the issue, such a tacit agreement could hardly have been made. Since tactical weapons are small and mobile, however, information on compliance will still be important.

Each side will presumably verify compliance through NTM, but without a written agreement no special access has been provided and no cooperative measures have been formally required. Under the circumstances, though, one would predict that each side would voluntarily act as transparently as possible when demobilizing its tactical weapons, refrain from interfering with NTM, and provide the other with ample information demonstrating compliance, so that ambiguities and misperceptions do not threaten the cooperative relationship. If voluntary assurance measures are successful in this setting, they could be used to support other tacit agreements.<sup>233</sup> Given the number of complex agreements already in force, greater reliance on assurance would be a desirable way to contain costs.<sup>234</sup>

Some specific issues likely to arise in future negotiations also lend themselves to the use of assurance.<sup>235</sup> START II, for example, will ban

230. President Bush and President Yeltsin reached agreement on the outline of this treaty in June 1992. See Michael Wines, *Summit in Washington; Bush and Yeltsin Agree to Cut Long-Range Atomic Warheads; Scrap Key Land-Based Missiles*, N.Y. TIMES, June 17, 1992, at A1.

231. See Krass, *supra* note 94.

232. See Andrew Rosenthal, *U.S. to Give Up Short-Range Nuclear Arms; Bush Seeks Soviet Cuts and Further Talks*, N.Y. TIMES, Sept. 28, 1991, at 1 (U.S. initiative); Serge Schmemmann, *Gorbachev Matches U.S. on Nuclear Cuts and Goes Further on Strategic Warheads*, N.Y. TIMES, Oct. 6, 1991, at 1.

233. For a recommendation that tacit agreement be used to achieve arms reductions, see Russell Hardin, *Contracts, Promises and Arms Control*, BULL. ATOMIC SCIENTISTS, Oct. 1985, at 14. But see Richard B. Bilder, *Formal Treaties and Tacit Agreements: An Exchange*, BULL. ATOMIC SCIENTISTS, Apr. 1985, at 51.

234. See Patricia B. McFate & Sidney N. Graybeal, *The Price for Effective Verification in an Era of Expanding Arms Control*, 500 ANNALS 73 (1988).

235. See Ivan Oelrich, *The Changing Rules of Arms Control Verification: Confidence Is Still Possible*, 14 INT'L SECURITY, Spring 1990, at 176. A recent incident demonstrates not only the value of the assurance approach, but the variety of mechanisms that can be used to convey assurances. President Yeltsin of Russia announced early in 1992 that the Soviet Union had for many years conducted a biological weapons (BW) program in violation of the 1972 Biological Weapons Convention, see *supra* note 111, and that all BW work was to be halted. After the United States expressed concern over the effectiveness of the ban, which was obviously difficult to verify, Yeltsin announced a variety of steps to demonstrate compliance. These included granting access for

multiple warheads on land-based missiles.<sup>236</sup> Information on compliance with this obligation might best be provided through the use of active transparency and assurance measures, such as the incorporation of FRODs to demonstrate single-warhead capacity, the affixing of radiation-detection "tags" to warheads, and the regular opening of missile silos, combined with external monitoring. Increasing concern over the control of nuclear warheads, deeper cuts in the number of warheads,<sup>237</sup> and the freezing of European nuclear arsenals, proposed by many as the next steps in strategic arms control,<sup>238</sup> might render feasible the use of unique identification tags and other identification systems designed to demonstrate compliance with numerical limits.<sup>239</sup>

The concluding Act of the Negotiation on Personnel Strength of Conventional Armed Forces in Europe (also known as CFE 1A), signed in July 1992 by the parties to the CFE Treaty, included only non-binding commitments to limit military personnel within the CFE Treaty area, perhaps because of the difficulties in monitoring such commitments.<sup>240</sup> In the future, however, extensive data exchanges, serial number identification systems, restrictions on movement and location, transparent maneuvers, and similar assurance devices, along with NTM and some on-site inspection, might be able to provide a sufficient level of confidence to support binding troop reductions.

Several multilateral arms control negotiations, including talks on chemical weapons,<sup>241</sup> as well as biological weapons and ballistic mis-

---

inspectors from the United States and Britain to civilian and military sites; an inquiry, with U.S. and British participation, into the operations of a scientific institute suspected of illicit BW research; closure of the military department that had supervised the unlawful activities; closure of the major Russian BW test site; and a promise of continuing reports on steps taken to implement the ban. Michael R. Gordon, *Russia and West Reach Accord on Monitoring Germ-Weapon Ban*, N.Y. TIMES, Sept. 15, 1992, at A6.

236. See Wines, *supra* note 230.

237. Under the START II agreement, the United States and Russia will be permitted to deploy no more than 3,000-3,500 strategic warheads by the year 2003. *Id.* Shortly after the START II agreement was announced, the Senate Foreign Relations Committee voted to direct the Bush Administration to negotiate verification mechanisms covering the remaining nuclear warheads themselves, as well as the missiles and launchers that were the subject of earlier verification systems. The Committee's concern was less Russian cheating than the possibility of warheads falling into the hands of rebel groups, terrorists or other unstable forces. Michael R. Gordon, *Senate Calls for Monitoring of the Warheads in Russia*, N.Y. TIMES, July 3, 1992, at A2. While intrusive on-site inspections could give some confidence as to the control of warheads, their small size and mobility will make assurance devices an essential element in a warhead control program.

238. See, e.g., Michael L. Gordon, *The Last Arms Accord?*, N.Y. TIMES, July 16, 1991, at A1, A9.

239. See generally Fetter & Garwin, *supra* note 189.

240. See DEP'T ST. DISPATCH, *supra* note 145, at 560-62. The agreement includes a modest assurance scheme, to be implemented during CFE on-site inspections. *Id.* at 561. For a general discussion of the difficulties, see Robert D. Blackwell, *Conceptual Problems of Conventional Arms Control*, 12 INT'L SECURITY 28 (1988).

241. As this Article was being prepared for publication, in September 1992, the United Nations Committee on Disarmament reached agreement on a multilateral

siles, will also be pursued in the coming years. The United States has sometimes talked of a comprehensive arms limitation regime for the Middle East, involving chemical, biological and nuclear weapons as well as ballistic missiles.<sup>242</sup> The United States also has hopes of strengthening the non-proliferation regime through agreements to restrict the export of advanced arms, CW precursors, missile technology and the like. In all of these areas, innovative informational provisions may be required.

For one thing, the parties to the negotiations in these areas will have widely differing verification capabilities; as a result, smaller, less advanced states may be reluctant to cooperate.<sup>243</sup> The two most likely solutions to this problem are the creation of IVOs required to share data from monitoring, as proposed for the CW convention,<sup>244</sup> and the sharing of data from national monitoring efforts, as agreed upon in Open Skies. Even modest assurance devices, though, would also help offset disparities in monitoring capacity and increase the chances for broad multilateral agreements.

A second, more serious problem is that the parties to multilateral arms control agreements, even the advanced countries likely to agree on new export controls, will vary considerably in their level of commitment to the cooperative goals of the agreements. Both verification and assurance techniques will be necessary, but an emphasis on one or the other will be more appropriate for particular countries.

Two general possibilities might be considered. First, agreements in these controversial areas might provide for both kinds of informational devices. In the case of relatively trustworthy countries, the parties could rely primarily on assurance mechanisms, reducing the economic and political costs of the agreement. In the case of other countries, verification techniques could be utilized more aggressively. This approach resembles the practice of many administrative agencies in applying aggressive enforcement methods against firms with a reputation for trying to evade regulation, while relying on the voluntary compliance of firms known to be law-abiding.<sup>245</sup>

Second, assurance devices could be used as litmus tests. Negotiators could propose the adoption of detailed assurance provisions requir-

---

chemical weapons convention, following 20 years of negotiation. The agreement is expected to be approved by the U.N. General Assembly and opened for signature by early 1993. Michael R. Gordon, *Negotiators Agree on Accord to Ban Chemical Weapons*, N.Y. TIMES, Sept. 2, 1992, at A1. See also Gordon, *infra* note 242. The United States and Russia jointly pressed for conclusion of the negotiations, using their own bilateral agreement as a model. See *supra* text accompanying note 165.

242. See Michael R. Gordon, *U.S. Weighs a Plan for Arms Control in the Middle East*, N.Y. TIMES, May 14, 1991, at A1.

243. See Oelrich, *supra* note 235, at 182.

244. See *id.* at 183.

245. The interaction between agency and firms can be interpreted as a PD. See John T. Scholz, *Cooperation, Deterrence and the Ecology of the Regulatory Environment*, 18 LAW & SOC. REV. 179 (1984).

ing states to credibly demonstrate their compliance. For the states that agreed, such measures, combined with some external monitoring, might well be sufficient. States that did not agree, however, would essentially be signalling their interest in possible defection. Further negotiations with this group would have to focus on strict verification measures.