International Law and International Relations

Edited by Beth A. Simmons and Richard H. Steinberg

CAMBRIDGE www.cambridge.org/9780521861861

hinted at trade sanctions, and the conference of the parties of CITES threatened not to schedule its next meeting in Kyoto if Japan remained out of compliance.

If there are no objective standards by which to recognize an "acceptable level of compliance," it may be possible at least to identify some general types of situations that might actuate the deployment of political power in the interest of greater compliance. First, states committed to the treaty regime may sense that a tipping point is close, so that enhanced compliance would be necessary for regime preservation. As noted above, the actions against Japan on the ivory import ban may have been of this character. After the high visibility given to the CITES moves to ban the ivory trade, there would not have been much left of the regime if Japan had been permitted to import with impunity.

Second, states committed to a level of compliance higher than that acceptable to the generality of the parties may seek to ratchet up the standard. The Netherlands often seems to play the role of "leader" in European environmental affairs both in the North Sea and Baltic Sea regimes and in LRTAP.[66] Similarly, the United States may be a "leader" for improving compliance with the NPT, where its position is far stronger than that of its allies.

Finally, campaigning to improve a compliance level that states concerned would just as soon leave alone is a characteristic activity for NGOs, especially in the fields of the environment and of human rights. NGOs increasingly have direct access to the political process both within the treaty organizations and in the societies of which they are a part. Their technical, organizational, and lobbying skills are an independent resource for enhanced compliance at both levels of the two-level game.

CONCLUSION

The foregoing discussion reflects a view of noncompliance as a deviant rather than an expected behavior, and as endemic rather than deliberate. This in turn leads to de-emphasis of formal enforcement measures and even, to a degree, of coercive informal sanctions, except in egregious cases. It shifts attention to sources of noncompliance that can be managed by routine international political and managerial processes. Thus, the improvement of dispute resolution procedures goes to the problem

[66] See Peter M. Haas, "Protecting the Baltic and North Seas," in Haas, Keohane, and Levy, *Institutions for the Earth*.

of ambiguity; technical and financial assistance may help cure the capacity deficit; and transparency will make it likelier that, over time, national policy decisions are brought increasingly into line with agreed international standards.

These approaches merge in the process of "jawboning" an effort to persuade the miscreant to change its ways that is the characteristic form of international enforcement activity. This process exploits the practical necessity for the putative offender to give reasons and justifications for suspect conduct. These reasons and justifications are reviewed and critiqued in a variety of venues, public and private, formal and informal. The tendency is to winnow out reasonably justifiable or unintended failures to fulfill commitments that comport with a good-faith compliance standard and to identify and isolate the few cases of egregious and willful violation. By systematically addressing and eliminating all mitigating circumstances that might possibly be advanced, this process can ultimately demonstrate that what may at first have seemed like ambiguous conduct is a black-and-white case of deliberate violation. The offending state is left with a stark choice between conforming to the rule as defined and applied in the particular circumstances or openly flouting its obligation. This turns out to be a very uncomfortable position for even a powerful state. One example is the now demonstrated Iraqi retreat in showdowns with the UN-IAEA inspection teams.[67]

Enforcement through these interacting measures of assistance and persuasion is less costly and intrusive and is certainly less dramatic than coercive sanctions, the easy and usual policy elixir for noncompliance. It has the further virtue that it is adapted to the needs and capacities of the contemporary international system.

[67] For an account of the Iraqi response, see Sean Cote, *A Narrative of the Implementation of Section C of UN Security Council Resolution 687*.

5

Is the Good News About Compliance Good News About Cooperation?

George W. Downs, David M. Rocke, and
Peter N. Barsoom

In the past few years many social scientists interested in cooperation have turned their attention to the problem of compliance in international regulatory regimes. Much of the empirical research in this area has been conducted by a group composed mainly of qualitative political scientists and scholars interested in international law.[1] Its message is that (1) compliance is generally quite good; (2) this high level of compliance has been achieved with little attention to enforcement; (3) those compliance problems that do exist are best addressed as management rather than enforcement problems; and (4) the management rather than the enforcement approach holds the key to the evolution of future regulatory cooperation in the international system. As Oran Young notes, "A new understanding of the bases of compliance – one that treats compliance as a management problem rather than an enforcement problem and that has profound practical as well as theoretical implications – is making itself felt among students of international relations."[2] In short, not only are the dreary expectations born of factors such as relative gains concerns, collective action problems, anarchy, and fears of self-interested

[1] For example, see Arora and Cason 1995; Chayes and Chayes 1990; 1991; 1993a; 1993b; Duffy 1988; Haas, Keohane, and Levy 1993; Hawkins 1984; Mitchell 1993; 1994a; 1994b; 1995; Scholz 1984; Sparrow 1994; Young 1989; and 1994.

[2] Young's quotation is taken from the dust jacket of Mitchell 1994a.

An earlier version of this article was presented at the annual meeting of the International Studies Association, Chicago, February 1994. The authors thank Abram Chayes, Robert Keohane, Marc Levy, Ron Mitchell, Ken Oye, Michael Ross, the editor of *International Organization,* and the anonymous referees for their helpful comments. The authors also acknowledge the support of the John D. and Catherine T. MacArthur Foundation to the Center of International Studies, Princeton University.

exploitation incorrect but also the enforcement limitations that always have appeared to sharply bound the contributions of international law and many international institutions now appear to have been exaggerated.

In this essay we will argue that the empirical findings of this group, which we refer to as the "managerial" school, are interesting and important but that its policy inferences are dangerously contaminated by selection problems. If we restrict our attention to those regulatory treaties that prescribe reductions in a collectively dysfunctional behavior (e.g., tariffs, arms increases), evidence suggests that the high level of compliance and the marginality of enforcement result from the fact that most treaties require states to make only modest departures from what they would have done in the absence of an agreement. This creates a situation where states often are presented with negligible benefits for even unpunished defections; hence the amount of enforcement needed to maintain cooperation is modest. Nothing is wrong with this situation in itself, but it is unlikely to provide the model for the future that the managerialists claim. Even if we assume that the absolute value of the benefits generated by this small amount of regulation is relatively large, further progress in international regulatory cooperation will almost certainly require the creation of agreements that present far greater incentives to defect than those currently in place (e.g., more demanding environmental standards, fewer nontariff barriers, steeper arms reductions). We have precious little evidence that such progress can be obtained in the absence of better enforcement.

After discussing the problems posed by endogeneity and selection, we present the theoretical argument for linking enforcement level to what we call "depth of cooperation" and examine the extent to which deep cooperation has been achieved without enforcement. We then present a number of prominent exceptions to the managerial school's unqualified generalizations about the causes and cures of noncompliance. Finally, we discuss the strategic implications of the evolution of increasingly cooperative regimes.

THE MANAGERIAL THESIS

The bedrock of the managerial school is the finding that state compliance with international agreements is generally quite good and that enforcement has played little or no role in achieving and maintaining that record. In Abram Chayes and Antonia Chayes's words, what ensures compliance is not the threat of punishment but "a plastic process of interaction among the parties concerned in which the effort is to reestablish, in the microcontext of the particular dispute, the balance of advantage that

brought the agreement into existence."[3] For the members of the managerial school, "noncompliance is not necessarily, perhaps not even usually, the result of deliberate defiance of the legal standard."[4] On those rare occasions when compliance problems do occur they should not be viewed as violations or self-interested attempts at exploitation, but as isolated administrative breakdowns. The causes of noncompliance are to be found in (1) the ambiguity and indeterminacy of treaties, (2) the capacity limitations of states, and (3) uncontrollable social or economic changes.[5]

Not surprisingly, the managerial school takes a dim view of formal and even informal enforcement measures. Punishment not only is inappropriate given the absence of any exploitative intent but it is too costly, too political, and too coercive. As Ronald Mitchell notes, "Retaliatory non-compliance often proves unlikely because the costs of any individual violation may not warrant a response and it cannot be specifically targeted, imposing costs on those that have consistently complied without hurting the targeted violator enough to change its behavior."[6] As a result, according to Young, "arrangements featuring enforcement as a means of eliciting compliance are not of much use in international society."[7] Since sanctions usually are more successful against economically vulnerable and politically weak countries and "unilateral sanctions can be imposed only by the major powers, their legitimacy as a device for treaty enforcement is deeply suspect," as Chayes and Chayes point out.[8] ***

Instances of apparent noncompliance are problems to be solved, rather than violations that have to be punished. According to Chayes and Chayes, "As in other managerial situations, the dominant atmosphere is that of actors engaged in a cooperative venture, in which performance that seems for some reason unsatisfactory represents a problem to be solved by mutual consultation and analysis, rather than an offense to be punished. Persuasion and argument are the principal motors of this process."[9] The strategies necessary to induce compliance and maintain cooperation involve: (1) improving dispute resolution procedures, (2) technical and financial assistance, and (3) increasing transparency. The last is especially important: "For a party deliberately contemplating violation, the high

[3] Chayes and Chayes 1991, 303.
[4] Ibid., 301.
[5] Chayes and Chayes 1993b, 188.
[6] Mitchell 1993, 330.
[7] Young 1994, 74 and 134.
[8] Chayes and Chayes 1993a, 29.
[9] Chayes and Chayes 1991, 303.

probability of discovery reduces the expected benefits rather than increasing the costs and would thus deter violation regardless of the prospect of sanctions."[10]

THE ENDOGENEITY AND SELECTION PROBLEMS

It is not difficult to appreciate why the findings of the managerial school suggest that both international institutions and even international law have a far brighter future than most international relations specialists have believed for the past fifty years. Apart from sharply contradicting the pessimistic expectations of many realists and neorealists about the inability of cooperation and self-regulation to flourish in an anarchic world, they also run counter to the claims of cooperation researchers in the rational-choice tradition. Such researchers emphasize the centrality of enforcement concerns in regulatory environments and characterize them as mixed-motive games, where the danger of self-interested exploitation is significant, as opposed to coordination games, where it is not.[11] Such findings certainly add credibility to the frequent speculation that the rational-choice tradition's affection for the repeated prisoners' dilemma has led it to overemphasize enforcement and underemphasize the potential for voluntary compliance and noncoercive dispute resolution.

* * *

To even begin to overcome the problems that endogeneity poses for understanding the role of enforcement in regulatory compliance, we need to control for the basis of state selection; that is, those characteristics of international agreements that play the same role for states as musical difficulty does for the school orchestras. One likely candidate is what we have termed the depth of cooperation. International political economists define the depth of an agreement by the extent to which it requires behind-the-border integration with regard to social and environmental standards as well as with regard to the reduction of barriers to trade. Here, however, the depth of an agreement refers to the extent to which it captures the collective benefits that are available through perfect cooperation in one particular policy area. Given the difficulties involved in identifying the cooperative potential of an ideal treaty, it is most useful to think of

[10] Chayes and Chayes 1993a, 18.

[11] See, for example, Abreu 1988; Abreu, Pearce, and Stacchetti 1986; 1989; Bayard and Elliott 1994; Downs and Rocke 1995; Hungerford 1991; Martin 1992; Staiger 1995; and Sykes 1990.

a treaty's depth of cooperation as the extent to which it requires states to depart from what they would have done in its absence. If we are examining the critical subset of regulatory treaties that require states to reduce some collectively dysfunctional behavior like tariffs or pollution, a treaty's theoretical depth of cooperation would refer to the reduction it required relative to a counterfactual estimate of the tariff or pollution level that would exist in the absence of a treaty. Of course, the depth of cooperation that a treaty actually achieved might be quite different than this figure. Here we measure depth of cooperation by the treaty level because that is the figure which serves as the basis for judging the level of compliance. In the absence of a trustworthy theoretical estimate of this counterfactual, it could be based on the status quo at the time an agreement was signed or on a prediction derived from the year-to-year change rate prior to that time.

Either estimate of depth of cooperation is obviously quite crude. There are doubtless policy areas in which, for any number of reasons, the potential for cooperation is much smaller than others. In such cases our depth measure will make cooperation in these areas appear shallower than it really is. Yet if one is willing to concede, as both managerialists and more conventional institutionalists argue, that there are substantial cooperative benefits that are as yet unrealized in the areas of arms control, trade, and environmental regulation, this depth of cooperation measure provides a rough idea of what states have accomplished. We can in turn use it to interpret compliance data and help assess the role of enforcement. While this measure of depth is hardly perfect, there is no reason to expect that it is biased in such a way as to distort the relationship between the depth of cooperation represented by a given treaty, the nature of the game that underlies it, and the amount of enforcement needed to maintain it.

Depth of cooperation is important to track because just as the role of enforcement differs in mixed-motive and coordination games, it also varies within mixed-motive games according to depth. To appreciate the connection, consider the following model. States A and B are playing a repeated bilateral trade game in which each state in each period chooses a level of protection $P \in [0, \infty)$ that influences the level of trade. The utility of state A is denoted as $U_A(P^A,P^B)$, and the utility of state B is denoted as $U_B(P^A,P^B)$. We do not specify the functional form of these utilities but instead adopt a series of plausible assumptions detailed in Appendix A.[12]

[12] These assumptions also contain conditions on the response functions $R_A(P_B)$ and $R_B(P_A)$, which denote the optimal single-period response of one state to a particular level of protection (e.g., tariff) chosen by the other state.

We will adopt the convention of representing the trade game as a prisoners' dilemma. While some have argued that this pattern of incentives emerges from a variety of plausible circumstances, we assume it has emerged from electoral and financial incentives provided by interest groups working to protect domestic products from foreign competition.[13] If we consider only two particular levels of tariffs $P^A < P^A_\circ$ and $P^B < P^B_\circ$, then the four outcomes represented by each side choosing P or $P_\circ$ form a payoff matrix of the prisoners' dilemma type. In this case, each side prefers higher tariffs regardless of the choice of the other side, but both sides prefer mutual cooperation to mutual defection. Unlike the repeated prisoners' dilemma, the choices defined by the present model are continuous rather than discrete. Treaties can be set at any level below the noncooperative tariff rates. Cheating can be limited or flagrant. And punishments can range from a barely perceptible increase in tariffs that lasts for one period to a multiple of current tariffs that lasts indefinitely.

Under the assumptions of our model, if tariff levels are high, both states have an opportunity to benefit by devising an agreement to lower them. Nevertheless, there is an incentive to exploit the other party's trust; that is, A's optimal one-period response to side B's cooperative tariff level will always be to raise tariffs. Self-interest will prevent such cheating only if the consequences of cheating are greater than the benefits. To achieve a situation where this disincentive exists, states must resort to a punishment for defection. In this case, one punishment strategy prescribes that state A begin by observing the treaty, but if B violates it, even modestly, state A should respond by abrogating the agreement (or otherwise reducing its level of compliance) for some specified period of time. During cooperative periods each side's tariff is supposed to be limited to $\bar{P}^A < P^A_\circ$ and $\bar{P}^B < P^B_\circ$, while in the punishment periods both sides raise tariffs to some noncooperative level. The most extreme punishment strategy, often called the "grim strategy," occurs when the response to any violation is permanent reversion to the noncooperative Nash equilibrium. A punishment strategy is sufficient to enforce a treaty when each side knows that if it cheats it will suffer enough from the punishment that the net benefit will not be positive.

To make this more concrete, consider an example where the noncooperative tariff is at a level of 100 percent for each side, and plausible

[13] For the former argument, see Staiger 1995, 27. For the latter, see Grossman and Helpman 1994.

treaties would provide for symmetric reductions in tariffs for each side.[14] Figure 5.1 compares the one-period utility of both sides observing the treaty with the temptation to defect. The temptation to cheat in this model rises rapidly with the cooperativeness of the treaty, while the treaty benefits rise less rapidly. This is what imposes a limit on which treaties can be supported. Figure 5.2 shows the punishment periods necessary to support treaties of various sizes. A shorter period would make the treaty vulnerable to cheating because it would be insufficient to remove all of the gains from violating the treaty. For example, a treaty that specifies a 5 percent reduction in tariffs only requires a punishment of two periods; the best treaty that can be supported with the maximal punishment of infinite duration is 37.19 percent. The increase in the ratio of the benefit of cheating to the benefit of cooperating means that increasingly severe punishments are necessary to deter defection – here severity means length of punishment – as the benefits of the treaty and corresponding restrictiveness of its requirements increase. Although the rate of increase in utility with the increase in punishment length decreases, the utility obtainable by very long punishments is still many times that of the utility obtainable with punishment lengths of one or two periods. The essential point the graph demonstrates is the deeper the agreement is, the greater the punishments required to support it.

The only relevant criterion is that punishment must hurt the transgressor state at least as much as that state could gain by the violation. This does not imply that, say, a certain amount of trade restriction should be punished by an equal trade restriction (tit-for-tat); nor does it mean that the transgressor be punished at least as much as the transgressor's violation hurt the other party. Although both of these standards possess aspects of fairness, neither is relevant to supporting the treaty equilibrium. Fairness and justice must take a back seat to the correct disincentive.

The specific mechanism by which states punish violations is less relevant to the relationship between depth of cooperation and enforcement than is the magnitude of enforcement. Although we motivate the model by using a case of centralized enforcement for convenience, nothing in the analysis precludes effective decentralized enforcement schemes. Enforcement can occur through linkages, as in the case of the Soviet Union and United States during the Kissinger years; through formal institutions such as the

[14] Of course, in the multiperiod model, the feasibility of maintaining this treaty depends on the discount factor, δ, as well as on the previous parameters. In this case, we use a discount factor of $\delta = .95$, corresponding to an interest rate of 5 percent.

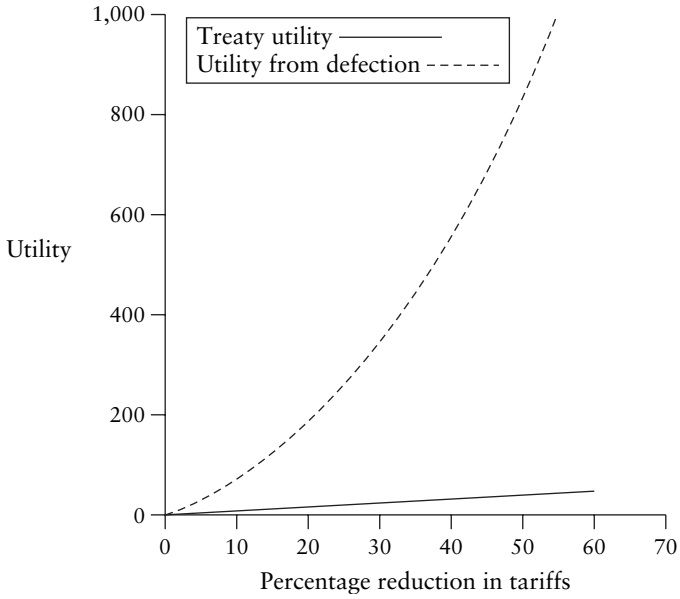

FIGURE 5.1. One-period utility of treaty compliance versus defection.

General Agreement on Tariffs and Trade (GATT) Dispute Settlement Procedure; through unilateral actions, as in the U.S. enforcement of fishery and wildlife agreements under the Pelly and Packwood–Magnuson amendments; or by domestic law as in the European Union and environmental treaties. Given the weakness of current international institutions and the relative difficulty in mobilizing formal sanctions, we suspect – like the majority of managerialists – that the most effective enforcement schemes may well be decentralized and not involve perfectly coordinated action by every signatory of a multilateral agreement.[15] This, however, does not negate the connection between depth of cooperation and the magnitude of the punishment necessary to maintain compliance in mixed-motive games.

DISCUSSION

This logical connection between the depth of cooperation represented by a given treaty and the amount of enforcement that is needed in mixed-motive games suggests that evaluating the importance of enforcement by examining how high compliance is when it is low or absent might

[15] On the role of decentralized enforcement schemes, see Ostrom 1990; and Kandori 1992.

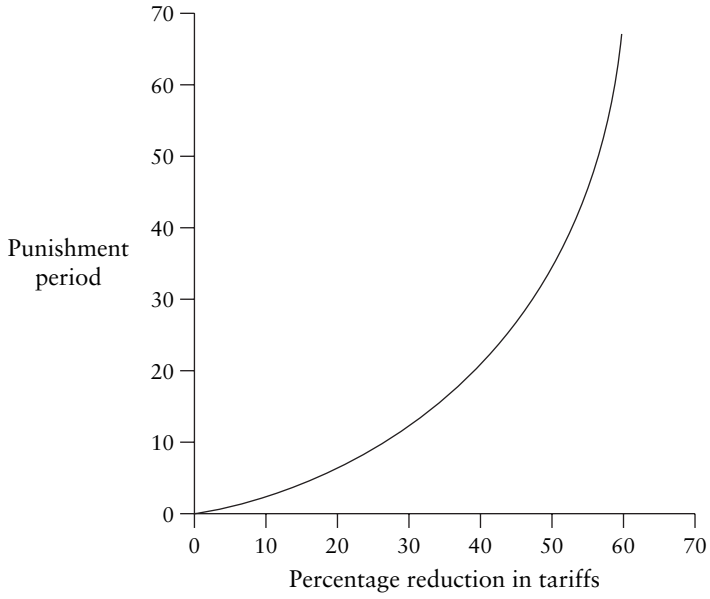

FIGURE 5.2. Punishment required to support treaties of various sizes

be misleading. We need to worry about the possibility that both the high rate of compliance and relative absence of enforcement threats are due not so much to the irrelevance of enforcement as to the fact that states are avoiding deep cooperation – and the benefits it holds whenever a prisoners' dilemma situation exists – because they are unwilling or unable to pay the costs of enforcement. If this were true, prescribing that states ignore enforcement in favor of other compliance strategies would be equivalent to telling the school orchestras to avoid wasting their time rehearsing. Just as the latter would condemn the orchestras to a repertoire of simple compositions, the prescriptions of the managerial school would condemn states to making agreements that represent solutions to coordination games and shallow prisoners' dilemmas.

* * *

Given the circumstances, it seems advisable to sidestep any attempt to inventory the nature of the underlying game and to evaluate some of the implications of the rival theories. We examine two. First, we will assess the depth of cooperation and the level of enforcement connected with prominent regulatory agreements that involve the reduction of behaviors that states have concluded are collectively counterproductive but that