# Preface

THIS BOOK starts from a very simple and unoriginal premise: actors who enter into a social interaction rarely emerge the same. More specifically, actors' behavior that prior to social interaction tended to diverge may converge as a result of this social interaction.[1] It tests the implications of this premise for cooperation in international relations by looking at an empirical puzzle: why would Chinese foreign policy decision makers—for the most part socialized in a relatively hard realpolitik strategic ideology, operating in an era of overwhelming and potentially threatening US power after the end of the cold war, and not offered obvious positive or negative material incentives—agree to cooperate in security institutions that did little to enhance China's relative power, and indeed had potential to do damage to its relative power interests?[2]

For mainstream international relations theories, this starting premise is at one and the same time an uncontroversial statement and a rather radical one. It is uncontroversial because mainstream IR accepts that social interaction can change behavior through the imposition of exogenous constraints created by this interaction. Thus, for instance, structural realists claim that the imperatives of maximizing security in an anarchical environment tend to compel most states most of the time to balance against rising power. Contractual institutionalists also accept that social interaction inside institutions can change the behavior of diverse actors in cooperative directions (e.g., changed strategies) by altering cost-benefit analyses as different institutional rules act on fixed preferences.

It is a radical starting point (at least for mainstream IR theory, but not for political science in general) if one claims that the behavior of actors

---

[1] By "convergence" I do not mean cooperation. I mean "increasing similarities" in basic characteristics of action. Actors' behavior can become increasingly similar and increasingly conflictual (as neorealists argue happens through "selection" in an anarchical environment). Nor do I use the term, as Botcheva and Martin do, to mean the isomorphism of state public policies as a function of participation in international institutions (Botcheva and Martin 2001).

[2] Of course, this is a puzzle only if one starts with standard realist assumptions that realpolitik actors operating under conditions of anarchy and facing more powerful states should be worried above all about preserving and enhancing their relative military power. It is not a puzzle if one does not start with this assumption. There should be, by now, plenty of skepticism in IR about the empirical and theoretical validity of this starting assumption, but the fact is that these assumptions still tend to define what the field considers "normal" and thus what the field considers puzzling. So this is where I start.

converges because of endogenous change in the normative characteristics and identities of the actors, or because of social identity–based, nonmaterial desires to conform. Put differently, convergence in the behavior of the participants in a social interaction may often have little to do with exogenous constraints and a lot to do with socialization (Wendt 1994: 384). This is, essentially, the claim made by those promoting the "sociological turn" in IR theory.

This book, then, is about socialization, and it is about whether socialization helps explain China's cooperation in major security institutions in the 1990s that had a potentially constraining effect on its relative power. More to the point, it proceeds from the constructivist claim that social interaction in international relations can affect actor interests in such a way as to then change the fundamental characteristics of the normative structures that constitute the world political system. Even more to the point, it looks at a critical but understudied link in this claim, namely the link between the presence of particular normative structures at the international level and the constraining effect of these norms on the behavior by the actor/agent at the unit level (whether the state or non-state actor level). Specifically, the book explores three microprocesses—mimicking, persuasion, and social influence—and examines how these work in cases of China's participation in a selection of international security institutions such as the United Nations' Conference on Disarmament in Geneva, the ASEAN Regional Forum and associated regional multilateral security dialogues, the Comprehensive Test Ban Treaty, and the antipersonnel landmine regimes.

The conclusion of the book is, in brief, that there is considerable, if subtle, evidence of the socialization of Chinese diplomats, strategists, and analysts in certain counter-realpolitik norms and practices as a result of participation in these institutions. Together with my earlier arguments about the origins of realpolitik norms and practices in Chinese history (Johnston 1995, 1996), I believe that this evidence casts doubt on a materialist explanation for realpolitik norms and practices rooted in the effects of international anarchy.

Constructivism, as Jeff Checkel has rightly point out, has been enamored by the sociological institutionalists' claim about the isomorphism in the norms and practices in world, regional, and local politics, but has not been very successful in explaining the microprocesses (Checkel 1998; Finnemore 1996a). For an approach whose central causal process is socialization, constructivist research thus far has been relatively quiet about how precisely actors are exposed to, receive, process, and then act upon the normative arguments that predominate in particular social environments. Yet, the ontology of social constructivism should point researchers squarely in the direction of these microprocesses: the susceptibility of

structures to minor perturbations, contingencies, nonlinearity, and path dependence set in motion by the conscious reflection and action of agents of change (e.g., ideas or norms entrepreneurs) all mean that it may indeed matter a lot, when explaining state behavior, how small groups, even individuals, are socialized through social interaction with other small groups and individuals in other states (and non-state entities).[3] Put simply, the value-added of the sociological turn in the international relations subfield rests heavily on showing what socialization microprocesses look like, how they are supposed to work, and whether they matter in a very common-sense way: whether they produce behavior—cooperative *or* conflictual— that would have been different had the actor not been exposed to the socialization processes in the first place.

To date, constructivism has only just begun to focus on socialization processes from the perspective of the "socializee." To be sure, this situation is changing with the emergence of some excellent detailed research on socialization processes in various European institutions. But it is still generally accurate to say that constructivist-oriented or -influenced research has not said much about how socialization works and why there are variations in the degree or completeness of socialization. And, of course, the new research that has looked at these questions in some detail has been limited to European institutions (and mostly not security institutions) where the empirical evidence needed for such detailed microprocess analysis is abundant relative to most other regions of the world.

Is there space for socialization arguments in IR theory? And if so, what would socialization arguments might have to look like to fill this space, that is, to ask questions central to the discipline that are unasked or badly or simplistically asked? Put differently, is there, implicitly or explicitly, a "demand" for socialization arguments? There is certainly empirical space created by other social sciences. There are healthy and robust research programs and cumulative empirical findings in social psychology, sociology, and political socialization on socialization processes of individuals and groups. But little of this has made it into IR theory.

There is also plenty of theoretical space in IR for socialization arguments. Realist theories say that socialization in the anarchical international environment is a key explanation for realpolitik practice. But realist socialization, in fact, is not socialization either in any standard social science definition or in any common use language sense; it is selection. It cannot account for the fact that many key actors are socialized into non-realpolitik practices and yet survive quite well under so-called anarchy. The democratic peace and democratic security community arguments are cases in point.

---

[3] For an excellent discussion of this ontology, see Cederman 1997.

Contractual institutionalism is right to point out that so much of state behavior goes on inside international institutions, which as Keohane even suggested, can serve as social environments (Keohane 1984). He comes to this conclusion from the observation that so much of a state's behavior can change through long-term involvement in institutions. But as contractual institutionalism has evolved, it primarily sees institutions as exogenous rule-based, sanction-based constraints on non-changing agents.

Socialization is central for constructivists, of course. Indeed it is, in a sense, the one process concept in IR that is uniquely constructivist.[4] But, as I will argue in chapter 1, constructivists have mainly focused on correlational analysis that suggests how ideas and identity "matter." In this respect, constructivism has been in a stage of development similar to contractual institutionalism a few years ago, keen to show theoretical plausibility, but less worried about how or why or to what degree ideational variables matter.

Finally, there is also a great deal of policy space for socialization arguments. After all, governmental and non-governmental diplomacy is often an effort to persuade, shame, cajole, and socially "pressure" states to change their collective minds and behavior. The concept of engagement, as assessed historically by Schweller, is at base one of a number of strategies toward rising major powers, in addition to balancing/containing, bandwagoning, and accommodating, but one aimed at changing the non–status quo elements of a rising power's strategy through the use of cooperative diplomatic measures (Schweller 1999). Certainly one can find this implied socialization assumption in the diplomacy of various states aimed at "engaging" China. One element of ASEAN engagement policy toward China has been an effort to "socialize" China into the rules of regional normative order (the so-called ASEAN way and its codes of conduct). One of the goals of Japan's ODA to China has been to turn China into a "responsible major power," "with self-awareness to contribute to the security and prosperity of the new international community after the end of the Cold War as well as of the East Asian region" (Kojima 2001).

---

[4] I do not want to leave the impression from this discussion that I believe various forms of realism, contractual institutionalism, and constructivism are alternative theories of cooperation. I do not consider constructivism a theory. Rather, like Wendt, I view it as an approach that can bundle a distinctive ideational ontology together with a wide range of extant "middle-range theories" about human and group behavior. Most fundamentally, like game theory, it lacks a basic foundational assumption about what motivates action. So this book is not about realism versus institutionalism versus constructivism as though these were three roughly equivalent theories of international relations. Rather it tests the plausibility of a set of explanations about cooperation, under the rubric of socialization, that are fundamentally different from arguments that derive from various versions of realism and contractual institutionalism.

The Clinton administration's strategy of constructive engagement was, for some, aimed at pulling China into the "international community," and exposing it to new norms of the market and domestic governance.[5] The policy was challenged by skeptics of engagement in Washington precisely because, in their view, it has failed to socialize China; that is, it failed to bring China into this putative "international community."[6]

But how precisely socialization is supposed to work through a diplomacy of engagement has never been all that clear, whether practiced by ASEAN, Japan, or the United States. It would seem worthwhile, then, to take seriously as a topic for academic inquiry a process that actual practitioners of international relations have believed is a reality in their world.

There is, then, lots of space to treat institutions as social environments. This means viewing microprocesses unique to social interaction that endogenously affect actor interest, preferences, and/or identities.

One useful way of exploring the phenomenon of socialization is through the study of China's participation in international institutions. Let me explain. The genesis of this study is, in part, my previous work on strategic culture (Johnston 1995, 1996b). It, too, reflected an effort to wrestle with the basic claim at the heart of mainstream IR theory, namely that anarchy and material power distributions are fundamentally determinative of the frequency and type of conflictual behavior in IR. My first book, then, was an effort to contribute to this debate, to see if indeed strategic cultures existed, if they did how would one know it, and if one could observe them, did they affect strategic behavior independently of anarchical material distributions of power under anarchy?

What I found, contrary to my initial expectations,[7] was that Chinese strategic culture in the Ming dynasty, as embodied in classical texts on

---

[5] See, for instance, the statements by various members of the Clinton administration: William Perry, "U.S. Strategy: Engage China, Not Contain," *Defense Issues* 10:109 at http://www.defenselink.mil/speeches/1995/s19951030-kaminski.html; Samuel Berger, "A Foreign Policy Agenda for the Second Term," Center for Strategic and International Studies, Washington, DC, March 27, 1997, at http://www.whitehouse.gov/WH/EOP/NSC/html/speeches/032797speech.html; Assistant Secretary of State Stanley O. Roth, "U.S.-China Relations on the Eve of the Summit," World Economic Forum, Hong Kong, October 14, 1997, at http://www.state.gov/www/policy_remarks/971014_roth_china.html; and Secretary of State Albright in "The U.S. and China," for *Diario Las Americas*, Miami, Florida, July 5, 1998, at http://secretary.state.gov/www/statements/1998/980705.html. Moderates in the Bush administration essentially endorsed a similar language about socializing China into a putative international community.

[6] For a representative example of the socialization discourse used even by critics of engagement, see Waldron 2001.

[7] The long-standing conventional wisdom is that China possessed a unique strategic culture that stressed defensiveness, anti-militarism, pacifism, and an official ideology of security based on magnanimity toward adversaries at peak points in relative power. The roots of

strategy and as manifested in strategy toward Mongol "threats," was a hard realpolitik one. In that period of history, Chinese strategic culture demonstrated a number of traits: the environment was considered to be highly conflictual; potential adversaries were described in zero-sum terms; offensive uses of force were considered highly efficacious, especially when relative power was favorable. And, there was an explicit axiom that strategy should flexibly respond to changing power circumstances free of political or moral restraints. There was indeed a strategic culture but a realpolitik one. Moreover, Ming strategy toward the Mongols reflected this realpolitik strategic culture. Ming strategists basically stressed that conflict with the Mongols was inevitable, that offensive strategies were best, and defensive or accommodationist strategies were useful only when relative power prevented the Ming from going on the offensive. Not surprisingly, then, the Ming rulers were more offensive strategically in the first seventy-five or so years of the dynasty. As power waned after this, strategies shifted to defensive and accommodationist ones. In short, as it turned out, the Ming read realpolitik texts, thought like realpoliticians, and acted like realpoliticians.[8]

Contrary to many readings of the book, it did not make an essentializing argument about a *Chinese* strategic culture across time. Nor was it an argument about the relationship between big *C* culture on the one hand and strategy on the other. It was emphatically not an argument about an inherent collective personality of the Chinese people. Rather it made an argument about the socialization of Chinese decision makers in particular periods of time (in this case, in the Ming dynasty) into a hard realpolitik strategic culture, a strategic culture that was and is not necessarily ethno-territorially bounded. As a socialization argument it held out the possibility that decision makers exposed to other strategic cultures could be socialized in alternative understandings of how to achieve security. In other words, it was an argument about strategic culture, not an argument about Culture and strategy.

Colleagues and critics who contended that material structures or material incentives and disincentives fundamentally structure the strategic choices of states read the book and said, in essence, "Thanks. You've developed procedures for extracting strategic culture from texts, analyzing effects on decision makers, and then analyzing effects on behavior, and in a society where the claims about the content of strategic culture

---

this were in early Chinese political philosophy, associated mainly with Confucian-Mencian thought: security was best achieved through beneficence, moral government, and enculturation, not force.

[8] In retrospect, given how the word "culture" tends to conjure up essentialized visions of groups, perhaps strategic ideology would have been a better term.

should lead to behaviors very different from dominant neorealist arguments, you come up with details about the ideational constructs that real-politicians develop. All you've done is to show in more detail how strategic culture is epiphenomenal to power realities of anarchy. End of story." And, some would add, "even if strategic culture is not epiphenomenal, when two hypotheses make the same predication, parsimony is the tie-breaker. And on these grounds, material structural explanations clearly are more parsimonious."

Needless to say, I am not yet willing to make these concessions. I have three reasons. First, there is no a priori reason to believe an ideational account of realpolitik is epiphenomenal to a material structural account. This has been a standard materialist realist assumption—that ideational variables may help explain deviant cases, but they cannot explain non-deviant cases. Yet logically, one does not follow from the other. There is no reason why ideational variables cannot also explain so-called normal or expected international behavior. To make this standard materialist realist claim, one would have to believe that there are arenas of human political and social activity that are "idea-less," or where perceptions/worldviews do not matter. This would require, then, a theory of why these are turned off when exogenous conditions are consistent with realist theory and turned on when they are not. I do not believe any version of realism has successfully developed such a theory.

Second, constructivists should be especially interested in making the ideational case for realpolitik. To this point, they have staked their claim to relevance by focusing mostly on what mainstream realist theories would call deviant cases or irrelevant cases (humanitarian intervention; weapons taboos; democratic norms in alliance cooperation; the development of European identity; norms of democratization or human rights; among others). But to question material realism and its variants requires reexamining cases and phenomena that materialist realism claims to explain. So there is no reason yet, until ideational arguments for realpolitik go nowhere,[9] for ceding realpolitik explanations to standard realist theories.

Finally, when two explanations make the same prediction, parsimony should not automatically be a tiebreaker. The parsimonious explanation may still be wrong. Rather the first response should be a critical test—spin out additional empirical implications that are competitive and see which set holds up.

---

[9] That is, arguments that treat realpolitik as a learned norm. See, for instance, Vasquez's (1993) claim along these lines. See also Rousseau's (2002 and 2006) very important work on the belief systems that underpin realpolitik preferences.

The present book, then, is broadly speaking a critical test: If realpolitik axioms embodied in realpolitik strategic cultures are epiphenomenal to anarchical structures, then they should not change as long as an anarchical material structure and its conflictual effects persists. If, on the other hand, realpolitik strategic cultures *are* independent of the anarchical materials structure, and are learned, absorbed for instance through exposure to key discourses and reinforced by experience, then they are, in principle, mutable or changeable. So, do realpolitik ideational structures change? How and under what conditions? And how would one test for this?

Here contemporary China provides a useful set of cases for exploring the plausibility of socialization arguments. Constructivists have posited that international institutions in particular are often agents of counter-realpolitik socialization. They suggest that there is a causal link between the presence of particular normative structures embodied in institutions and the incorporation of these norms in behavior by the actor/agent at the unit level. For one thing, the interaction with activists, so-called norms entrepreneurs, is most likely inside institutions. For another, social conformity pressures are more concentrated inside institutions. Third, inside institutions, interaction among agents on specialized issues and exchanges of specialized information are sustained and intense. Finally, institutions often have corporate identities, traits, missions, normative cores, and official discourses at odds with realpolitik axioms. So, for example, some arms control institutions expose actors to an ideology where interalia: multilateral transparency is better than unilateral non-transparency; disarming is better than arming as a basis of security; cooperative security is better than unilateral relative power strategies for achieving state security; and evidence of cooperative potential is greater than evidence of fixed conflictual environment. All of these axioms challenge aspects of realpolitik ideology.

So, if there is any counter-realpolitik socialization going on, it ought to be happening in international institutions. But to test this, ideally, one needs a state where the predominant security ideology prior to involvement in institutions is at a maximum distance from that of the institutions (e.g., hard realpolitik, unilateralist). One also needs a "novice" state (socialization, after all, often refers to the process of inducting newcomers such as youths, immigrants, recruits, new states) into the membership norms and practices of a social group. In particular, one needs a tabula rasa state that then becomes rapidly involved in international institutional life. China is such a state.[10] Its strategic elites since 1949 have been social-

---

[10] It is important to note that I usually use "China" as a shorthand for China's foreign policy decision makers or policy officials and representatives. As is clear from the main chapters, the unit of analysis is individuals and small groups. "China" is often more convenient than "Chinese officials with the most influence on foreign policy decision *X*."

ized in the hard realpolitik of Marxism-Leninism, modern Chinese nationalism, and, for Mao at least, elements of classical strategic thought. It has been one of the most dispute-prone major powers from 1949 to the 1980s (after the United States). Until 1980, it was essentially uninvolved in international institutions, arms control or otherwise. But by the late 1990s, its participation rates were not significantly different from other major powers, and, in comparison to its level of development, it was overinvolved.

The notion of using China's participation in international institutions in the late twentieth century as a critical test of my arguments developed from a study of Ming strategic culture was not, of course, the only inspiration for this project. Another was a dissatisfaction with the absence of details about how precisely one could actually observe the socialization processes that constructivism claims are continuously at work in world politics. I am not an international relations theory architect. I see myself as an engineer. I need to understand how one applies the theories and claims and arguments produced by constructivism's architects (Wendt, Ruggie, Finnemore, Katzenstein, Kratochwil, Meyer, Onuf, and Hopf, among others). How does one do empirical work if one is hoping to test constructivist arguments? When I began this project, there were not many engineers around.

I am relieved to find out just how many people shared my concerns about empirical processes.[11] In the past handful of years, a small number of detailed, process-oriented, richly descriptive, and inferential research on the conditions of socialization in international institutions have been published.[12] Some of this work has discovered processes and effects of social interaction inside institutions that accord with my findings in the Chinese case, as I will discuss in more detail later. However, almost all of this work has focused on European institutions. Scholars have focused either on the processes of socialization experienced by bureaucrats, managers inside various EU institutions, or on those experienced by novices to European institutions—the new entrants from various Eastern European countries, for example—as they are exposed to deliberate efforts by extant European institutions and their norms entrepreneurs to inculcate Eu-

---

[11] Jeff Checkel and Thomas Risse have been the most persistent and insightful in calling for greater attention to the empirical microprocesses of socialization.

[12] See, for instance, the special issue of *International Organization* (Fall 2005) on socialization in European institutions; the special issue of *Comparative Political Studies* (February–March 2003) on rationalist and constructivist approaches to understanding how the EU functions, especially Checkel 2003 and Lewis 2003; Hooghe's test of rationalist and constructivist explanations for EU officials' images of Europe (2002); Kelley's (2004) test of institutional conditionality and socialization explanations for change in minority policies in Eastern Europe; Schimmelfennig 2002; and Gheciu 2005.

ropean values and identity. Not that these are exactly easy cases or most likely cases, but for the most part, the objects of socialization—say, the bureaucrats moving from national socialization environments to international or supranational ones—express lower levels of resistance, and their material interests at stake are less frequently in the realm of "high" security politics than the China cases I examine in this book. On case selection grounds, then, testing hypotheses about socialization on Chinese participation in international security institutions predominantly in an era of unipolarity and on issues where material power interests are at stake can help strengthen the conclusions about socialization in this burgeoning literature on Europe.

This focus on China's behavior inside international (security) institutions is work built on the shoulders of many expert scholars in the field. Participation in international institutions has not been a major focus of Chinese foreign policy studies, despite the fact that some of the richest behavioral data about Chinese foreign security, economic, and cultural policy is found in the statements, votes, and behind-the-scenes interactions by Chinese diplomats inside these institutions. But there is a small and growing collection of very rich empirical research that is implicitly, sometimes explicitly, sensitive to socialization processes that can occur inside institutions.

Samuel Kim's pathbreaking body of work shows how from the 1970s to the present, China has moved from "system-transforming" to "system-reforming" to "system-maintaining" preferences inside the United Nations system. The persistence of free-riding behavior into the 1990s does not imply a fundamental desire to dismantle or radically overhaul the UN more to China's advantage. Not all of this shift in China's basic approach to UN institutions is a function of participation alone. As in much of China's diplomacy of the 1980s and 1990s, the driver has been domestically generated, a desire for rapid economic growth that benefits the legitimacy of Communist Party rule. But the system-maintaining preferences have also been anchored by multilayered material and ideational incentives and constraints, including diffuse image, that create new indifference curves linking interests that, in the past, were not connected (e.g., Kim 1999).

Harold Jacobson and Michel Oksenberg's pioneering study of China's participation in international economic institutions in the early 1980s showed how this led to organizational, ideational, and material responses in the Chinese foreign policy process that encouraged a deepening of China's engagement with the rest of the world. In some ways, their book anticipated some of the concepts of interest to constructivists. They described processes of change in Chinese diplomacy that are variants of mimicking, persuasion, and social influence. They note that participation

in institutions can lead to greater cooperation as new domestic policy institutions are created to mediate interaction with the international institution (this has some similarities with mimicking). Or it can lead to cooperation because of the intensive interaction, often inside the international institution itself, between agents of the state and agents of the institution. This interaction, in turn, can lead to learning and internalizing the institution's norms (this resonates with propositions about persuasion).[13] Or it can lead to cooperation because of the rewards and penalties offered by the institution (Jacobson and Oksenberg 1990:6).[14]

Likewise, Pearson shows that those in the Chinese policy process most extensively interacting with officials from the World Bank and IMF are those who are most committed to transparency in policy making and to exchange rate predictability (Pearson 1999:224). Cooperation has been elicited by a range of foreign policy goals: from concerns about maintaining an image as a "team player" to heading off the use of economic coercion against China to using international commitments as a lever against opponents to more fundamental domestic economic reform (Pearson 1999b).

Elizabeth Economy's work on China's participation in environmental institutions such as the Framework Convention on Climate Change process shows that, while overall conservative pro-development organizational interests eventually came to dominate the policy process, the repository of more proactive views on China's role in reducing greenhouse gases was among scientists in environmental research institutes who interacted most intensively with Western climate change specialists (Economy 1994).

Ann Kent's study of China's approach to international human rights regimes shows how, within the broad goal of minimizing threats to the rule of the Communist Party, China's diplomacy in this realm has been constrained in part by an acceptance of procedural norms of participation inside institutions, and by a sensitivity to multilateral praise and blame (Kent 1999).

The common thread in much of this work is that the persistence of realpolitik-derived concerns about preserving sovereignty and autonomy has been moderated in certain instances either by changes in definitions

---

[13] They show, for instance, how Chinese participation in the IMF and World Bank helped to spread Ricardian economic concepts inside the Chinese policy process (Jacobson and Oksenberg 1990:143–44).

[14] They did not differentiate between material and social incentives. For another excellent, more recent, study of how precisely participation in less high-profile institutions (educational, for instance), mediated by the nature of China's local bureaucratic political structures, created positive incentives for greater participation in global economic and cultural processes, see Zweig 2002.

of interest or by linking realpolitik interests to other values, such as image and status, in new policy trade-offs. There appears to be a common recognition across these works that China's socialization inside these institutions is a "work in progress"—fragile, and susceptible to the deeply ingrained hyper-sovereigntist crisis response mechanisms of a regime with shaky legitimacy.

As a general rule, however, the work on China's multilateral diplomacy has only just begun to tap into theories and methods offered by the international relations field. The concept of socialization tends to remain undefined, the separate microprocesses through which social processes can constrain behavior left unclear. There is, at times, an implicit assumption that socialization is teleological, that progress toward acceptance of an institution's pro-social norms is socialization, while regression is evidence for the absence of socialization.[15]

This book does not radically challenge the flavor of the arguments of this burgeoning literature on China's participation in institutions. Indeed, it is not surprising that a literature rooted in the rich analyses of one country—local knowledge, so to speak—has a more constructivist feel to it than more general IR treatments of cooperation in institutions. Area

---

[15] Often the causal arguments explaining variation in the quantity and quality of China's cooperation are not clearly differentiated, and the implications for broader disciplinary and even policy questions about the conditions for conflict and cooperation are left underaddressed. Moore and Yang suggest, for instance, that a hybrid concept of "adaptive learning" best accounts for evidence of both minimax cooperation and genuine preference change (Moore and Yang 2001:228). Admittedly, learning is a slippery term in IR theory. But it is also clear that, conceptually, adaptation and learning are fundamentally different microprocesses. To put it metaphorically, adaptation refers to tactical shifts in cooperation, say, by a player with prisoners' dilemma preferences, as the exogenously imposed relative costs of defection increase (say, through the offer of side payments or the threat of sanctions). Learning, however, is most usefully viewed as a change in the basic preferences of the player, a shift away from one type of preferences through intensive socialization processes (which could shift preferences in either more cooperative or more conflictual directions). Although in practice these causal processes may be sequenced, they need not be, and they need to be kept analytically separate because they have very different explanations for cooperation and for its durability. Adaptation in the context of a prisoners' dilemma game suggests that once exogenous incentives change, players will move back to the old pattern of opportunistic defection. Learning suggests that cooperation can be sustained even as the material incentives dissipate. Though I do not use the learning versus adaptation framework in this book, it roughly maps onto the differences between persuasion and social influence. As I point out in the book, the institutional designs conducive to these different processes are not the same. Thus, whether an actor's cooperation is a function of learning (persuasion) or adaptation (social influence) implies very different policy prescriptions about institutional design. For more details on my distinction between learning and adaptation, see Johnston 1996b:31–33. For a summary of the learning theory literature, see Levy 1994. For a discussion of the sequencing of strategic and nonstrategic causes of behavioral change, see Jupille, Caporaso, and Checkel 2003:22–23.

specialists can be more attuned to the important nuanced changes in policy discourses and outputs that are not easily coded into the typologies typical in the discipline. Many are predisposed to be sympathetic, if only implicitly, with a theoretical approach that ontologically leans toward historical contingency and intensive endogeneity in social and political processes.[16] Rather this book tries to refine, clarify, and test more systematically the social processes that might lie behind changing levels of cooperation, and to show how IR theory both illuminates and benefits from the intensive study of China's international relations, in particular China's policy toward international security institutions.[17]

The book begins with a review of the status of socialization in various clusters of theory in the international relations subfield today. It then outlines three microprocesses of socialization—mimicking, social influence, and persuasion.

Mimicking explains pro-group behavior as a function of borrowing the language, habits, and ways of acting as a safe, first reaction to a novel environment: "I will do $X$ because everyone seems to be doing $X$ and surviving. So, until I know better, $X$ is what I will do." But by doing $X$, the actor can fall into habits of discourse and practice that constrict its options down the road. It is not the same thing as a rational search for successful exemplars (what can be called emulation).

Social influence explains pro-group behavior as a function of an actor's sensitivity to status markers bestowed by a social group, and requires some common understanding in the social value the group places on largely symbolic backpatting and opprobrium signals: "I should do $X$ because others believe $X$ is the appropriate thing to do and I will be rewarded socially for doing so." The chapter then develops some simple hypotheses about how different institutional designs should be conducive to different kinds of socialization processes, and then justifies the research design choices I have made, noting why China's participation in international security institutions in the 1990s, primarily, is a good initial test of these three microprocesses.

Persuasion explains pro-group behavior as an effect of the internalization of fundamentally new causal understandings of an actor's environment, such that these new understandings are considered normal, given,

---

[16] But not all. There is a tradition in area studies of sweeping cross-temporal, primordialist generalizations about an area as well. Some of the most vigorous "orientalizing" has been done by area specialists. For a nice statement of the argument that good theory testing and development rests on good local or area knowledge, see Christensen 1998.

[17] Thus far, the vast majority of work on China in multilateral institutions has focused on economic and environmental institutions, not security institutions. For exceptions, see Johnston 1986, 1996b; Medeiros and Gill 2000; and Medeiros 2006.

and normatively correct: "I should do *X* because it is good and normal for me (us) to do so."

Chapters 2 through 4 then look separately at each of these micro-processes and some least likely cases that appear to offer some initial confirmation of the plausibility of various forms of socialization. They then outline some additional empirical implications that one might expect to see if the causal claims in the main cases are correct. The chapters move from most to least "rationalist" of these microprocesses.

Chapter 2 looks at mimicking and how the mere fact of having to participate in some of the first major international arms control institutions, such as the Conference on Disarmament in which Chinese diplomats participated, led the PRC to adopt in limited fashion the habits, language, and even organizational models of arms controllers in these institutions. Before a clear security "interest" in participation developed, the PRC found that in order to participate in any meaningful way, it had to adopt some of the more common forms of participation.

Chapter 3 looks at social influence, the impact of social backpatting and opprobrium on cooperation. The main cases are the Chinese decision to sign the CTBT and the Protocol II of the Conventional Weapons Convention. In both instances, the military and economic costs of signing were not trivial. In both cases, signing was not induced by material incentives or disincentives. In both cases, there is considerable evidence that diffuse image—the desire to minimize opprobrium costs—weighed heavily on Chinese leaders as they calculated the costs of these agreements.

Chapter 4 looks at evidence for persuasion as a social process that helps account for the development of a limited, protomultilateralist discourse and practice, especially among those most directly involved in regional security institutions such as the ASEAN Regional Forum. It finds that although some of the multilateralist discourse has been hijacked by those who see multilateral institutions as potential anti-US tools, the original sources of this discourse came from extensive and intensive interaction in regional institutions (the ARF and related processes) whose designs were conducive to persuasion.

The tests in these chapters are eclectic ones. They are designed to explore the plausibility of these three microprocesses under fairly tough case selection conditions, and then to explore other empirical extensions. The chapters can be read mostly as stand-alone chapters, so the reader is welcome to skip one or more or read them in any particular order. The point is not to test which of the three microprocesses best explains Chinese cooperation in international institutions. Rather the goal is to show how each works.

Before getting to the substantial part of the book, I want to note at the start what this book does not claim. First, it does not claim that through

socialization in international institutions *all* Chinese leaders and officials have fundamentally changed their strategic culture(s) or ideologies, and have abandoned realpolitik for notions, say, of cooperative security. As will be evident, I am making more limited claims—albeit within the context of hard or least likely cases—about how under certain conditions certain parts of the decision-making process have been weaned away from realpolitik calculations of maximizing relative power. Put differently, there is now greater tension within the PRC's overall diplomatic thinking and practice between harder realpolitik and softer idealpolitik than ever before. But I am not arguing that multilateralism, say, has supplanted realpolitik as the predominant ideational construct behind China's foreign policy.

Second, this is not a book about constructivism versus rationalism, where the "logics of appropriateness" trump the "logics of consequences" in all cases. As is evident, at least two of the socialization microprocesses I will discuss could fall within the "rationalist" paradigm (mimicking and social influence)—an actor is, roughly speaking, maximizing some utility by choosing alternatives that appear to increase the probability of meeting some goal. In the case of mimicking, survival is being maximized by copying the group. In the case of social influence, self-perceptions of status are being maximized through interaction with other humans. Where I part from most versions of rationalism, however, is that these rationalisms are social, not material. Only persuasion entails a process that might fall clearly within the rubric of the logics of appropriateness, where socialization leaves actors with new definitions of self that provide self-evident and normal notions of expected behavior. The reality is that socialization, broken down this way, does not fit neatly into either a constructivist or a rationalist approach.

# Persuasion

PERSUASION has to do with cognition and the active assessment of the content of a particular message. As a microprocess of socialization, it involves changing minds, opinions, and attitudes about causality and affect (identity) in the absence of overtly material or mental coercion. Or, in Perloff's words, persuasion is an "activity or process in which a communicator attempts to induce a change in the belief, attitude, or behavior of another person or group of persons through the transmission of a message in a context in which the persuadee has some degree of free choice" (Perloff 1993:14).[1] It can lead to common knowledge, or "epistemic conventions" (that may or may not be cooperative), or it can lead to a homogenization of interests. That is, actors can be persuaded that they are indeed in competition with each other, or that they share many cooperative interests. The point is, however, that the gap or distance between actors' basic causal understandings closes as a result of successful persuasion.

Persuasion is a prevalent tool of social influence. Social psychologists have shown, for instance, that in interpersonal relationships people tend to rank changing others' opinions very high in a list of influence strategies, regardless whether the other is considered a friend or an enemy (Rule and Bisanz 1987:192). Some political scientists have called persuasion the "core" of politics, the "central aim of political interaction" (Mutz, Sniderman, and Brody 1996:1). In Gibson's view, politics is all about persuasion: "Real politics involves arguments; it involves people drawing a conclusion, being exposed to countervailing ideas, changing views, drawing new conclusions" (Gibson 1998:821). Communications theorists have argued that all social interaction involves communications that alter people's "perceptions, attitudes, beliefs and motivations" (Berger 1995:1). How persuasion works, therefore, is a focus of a great deal of research in communications theory, social psychology, and sociology, and there is no obvious way of summarizing a disparate and complex literature (see Zimbardo and Leippe 1991: 127–67).[2] But let me try.

---

[1] I would eliminate "or behavior" from this definition, as presumably behavior cannot change in the face of persuasion attempts without a change in some belief or attitude about the validity of the behavior.

[2] Despite the volume of this literature, "[t]o date there is precious little evidence specifying who can be talked out of what beliefs, and under what conditions" (Berger 1995:8).

Essentially, an actor is persuaded in three ways. First, s/he can engage in a high-intensity process of cognition, reflection, and argument about the content of new information (what Bar-Tal and Saxe call cognitive capacity, 1990:122). Also known by some as the "central route" to persuasion, the actor weighs evidence, puzzles through counter-attitudinal arguments, and comes to conclusions different from those he/she began with. That is, the merits of the argument are persuasive, *given* internalized standards for evaluating truth claims. Arguments are more persuasive and more likely to affect behavior when they are considered systematically and, thus, linked to other attitudes and schema in a complex network of causal connections and cognitive cues (Wu and Shaffer 1987:687; Petty, Wegener, and Fabrigar 1997:616; Zimbardo and Leippe 1991:192–97). This process of cognition, linking one set of attitudes to another, is more likely to occur when the environment cues and allows for the actor to consider these connections. That is, it is less likely to be spontaneous than it is promoted. As Gibson has shown with political intolerance among Russian voters, levels of intolerance and tolerance toward political opponents, and the overall balance between these two extremes, will change if counter-attitudinal arguments are presented to respondents that compel them to "think harder" about the implications of their initial attitudes. Thinking harder simply means people are cued, and have the time, to connect the implications of their initial attitude to outcomes that might affect their interests based on different sets of attitudes. Thus, an initially intolerant view might change to a more tolerant one if the respondent is cued to think about the implications of cycles of intolerance for political stability or for opportunities for themselves to present their own political opinions in the face of opposition (Gibson 1998:826–31). Thus, the probability of some change in attitudes through cognition increases in an iterated, cognition-rich environment (where there is lots of new information that cues linkages to other attitudes and interests). As a general rule, then, the probability goes down if the initial attitudes are already linked to a larger, internally consistent "network of supportive beliefs," particularly if these beliefs are about a high-threat group, a "crystallization hypothesis" applied to potential enemies (Sniderman, Brody, and Tetlock 1991:210; Gibson 1998:844).

Second, the actor is persuaded because of her/his affect relationship to the persuader; sometimes called the "peripheral" route, here the persuadee looks for cues about the nature of this relationship to judge the legitimacy of counter-attitudinal arguments. Thus, information from in-groups is more convincing than that from out-groups. Information from culturally recognized authorities (e.g., scientists, doctors, religious leaders) is more convincing than that from less authoritative sources. This will be especially true for novices who have little information about an issue on

which to rely for guidance (Zimbardo and Leippe 1991:70).[3] Information from sources that are "liked" is more convincing than that from sources that are disliked. Liking will increase with more exposure, contact, and familiarity. The desire for social proofing means that information accepted through consensus or supermajority in a valued group will be more convincing than if the group were divided about how to interpret the message (Petty, Wegener, and Fabrigar 1997:612, 617, 623, 627, 629; Kuklinski and Hurley 1996:129–31; Napier and Gershenfeld 1987:159; Isen 1987:206, 210–11; Axsom, Yates, and Chaiken 1987:30–31).[4] A favorable endorsement of a candidate or idea in politics by others can lead an actor to adopt a more favorable opinion as well (Bartels, cited in Hirshleifer [1995:201]).[5]

Third, the persuasiveness of a message may be a function of characteristics of the persuadee her/himself. This can refer to a range of variables from cognitive processing abilities, to the strength of existing attitudes (usually these are stronger if developed through personal experience than if based on hearsay or indirect experience, for example), to what appears to be a deeply internalized desire to avoid appearing inconsistent, to the degree of independence an agent might have in relation to a principal. Thus, for example, an attitude associated with an explicit behavioral commitment made earlier will be more resistant to change later because actors experience discomfort at being viewed as hypocritical and inconsistent. Conversely, a new set of attitudes will be more persuasive if associated with a new, high-profile behavioral commitment (Cialdini 1984; Wu and Schaffer 1987:677). Thus, a focus on the characteristics of the persuadee means looking at the individual features that can either retard or propel persuasion. All this means is that actors entering a social interaction bring with them particular prior traits that, interacting with the features of the

[3] Or as Gibson puts it, "Especially when people do not have much experience with political institutions and processes, it is easy to imagine that their initial viewpoints are poorly anchored in a highly articulated and constrained belief system, and that considerable potential for effective persuasion exists" (Gibson 1998:821).

[4] Using different language, Habermasian constructivists make a similar point: "[T]rust in the authenticity of a speaker is a precondition for the persuasiveness of a moral argument" (Risse 1997:16; see also Williams 1997:291–92). See also Bourdieu 1991:109–11. Game theorists have come to a similar conclusion, only using different language. Lupia and McCubbins (1998) note that persuasiveness rests basically on the persuadee's belief that she or he shares common interests with the persuader and that the information the persuader is offering benefits both. They do not specify what kind of information leads to the first belief. But it could, in principle, be anything from the list in the previous paragraph.

[5] Crawford (2002:36) suggests that the emotional appeal of an argument adds to its persuasiveness. Arguably, the affective relationship between persuader and persuadee subsumes emotional cueing.

social environment and other actors, lead to variation in the degree of attitudinal change.[6]

Lupia and McCubbins (1998) argue that all of these conditions and characteristics are simply indicators of more basic conditions for persuasion, namely, that the persuadee believes the persuader to be knowledgeable about an issue and that his or her intentions are trustworthy. The more certain the persuadee is about these beliefs, the more likely the persuader will be persuasive. This certainty can be a function of familiarity and extensive interaction that, over time, reveal the persuader's knowledge and trustworthiness. Or it can be a function of "external forces" that make it difficult or costly for the persuader to hide knowledge (or lack thereof) and trustworthiness (e.g., mechanisms for revealing knowledge, penalties for lying, costly actions that reveal the position of the persuader). Any other factors, such as ideology, identity, culture, and so on, are only predictors of persuasion to the extent that they reveal information to the persuadee about the persuader's knowledge and trustworthiness.

Lupia and McCubbins's rigorous formal model of persuasion is probably correct in stripping the process down to these two pieces of perceived information. But they miss the more interesting question about the empirical frequency with which social variables such as perceived ideology, identity, and/or cultural values are in fact the primary cues that people use to determine the degree of knowledge and trustworthiness of a persuader. That is, on average is perceived shared identity between persuadee and persuader more likely to be used by the persuadee as an authoritative measure of a persuader's knowledge and trustworthiness than other kinds of cues? The answer has important implications for how social interactions lead to socialization and how different institutional designs might lead to different socialization paths. Lupia and McCubbins tend to focus, as befits their interest in signaling games, on the role of external forces in clarifying beliefs about knowledge and trustworthiness of persuaders. Since, they argue, social and political environments are rarely ones where persuader and persuadee interact face to face over long periods of time, the familiarity/personal interaction route to beliefs about the persuader's knowledge and trustworthiness tends to be less common.

---

[6] Of course, persuasion in practice is likely to be a combination of all these microprocesses. Jorgensen, Kock, and Rorbech found in a study of televised political debates in Denmark, for example, that the most persuasive debaters were those who used a small number of extended, weighty discussions of specific qualitative examples. The use of these specific, straightforward, and logical examples seemed to accentuate the authoritativeness of the debater and were easier for viewers to assess and adjudicate. See Jorgensen, Kock, and Rorbech 1998.

This may be true at the national level of persuasion (e.g., political messages from politicians aimed at masses of voters), but it is not necessarily true at the level of social interaction in international institutions among diplomats, specialists, and analysts. Here the first route—familiarity and iterated face-to-face social interaction—may be more common. Hence, affect based on identity, culture, and ideology may be more critical for persuasion than external forces and costly signals. Institutions, therefore, that are "lite" in terms of these external forces nonetheless may create conditions conducive to persuasion—and convergence around group norms—even though there are few material incentives for the persuader to deceive and few material costs for the persuadee to defect from the group. I will come back to this at the end of this chapter.

Obviously, persuasion in the end is a combination of all three processes just outlined, and it is hard to run controls that might isolate the effects of any one process. People are more likely to think hard and favorably about a proposition, for instance, when it comes from a high-affect source, in part because affect helps kick in resistances to information from other sources (Mohr 1996:81–82). On the other hand, one can identify ideal combinations that could, in principle, be tested. Given, then, an effort by a persuader to provide information with a view to changing basic principled, causal, or factual understandings, certain kinds of social environments, therefore, ought to be especially conducive to persuasion:[7]

- when the actor is highly cognitively motivated to analyze counterattitudinal information (e.g., a very novel environment);

- when the persuader is a highly authoritative member of a small, intimate, high-affect in-group to which the persuadee also belongs or wants to belong;

[7] For a similar list of hypotheses, see Checkel 2001:222. For a differently worded list that nonetheless can map onto a number of these hypotheses, see Crawford 2002:36. Moravcsik (2001) claims that because one can imagine "rationalist" arguments that make similar predictions, these kinds of "mid-range" theory hypotheses developed by constructivism are somehow subsumed by rationalist approaches, or are at the very least theoretically indistinguishable from so-called rationalist predictions about persuasion. This is debatable on a number of grounds. First, as I note in reference to the only systematic contractualist argument about persuasion (Lupia and McCubbins 1998), even they admit that there may be different, more affect-based reasons for persuasion in face-to-face interactions that are not captured by contractualist or "rationalist" microprocesses. Second, since the microprocesses in social psychological–derived hypotheses are different, the practical implications for the kinds of institutional designs most conducive to persuasion are meaningfully different. Finally, Moravcsik misses the point of critical tests—namely, they are set up precisely because two different sets of theoretical arguments make, in a specific instance, similar predications about behavior. The fact that they do prior to a critical test says nothing about which approach is distinctively superior.

- when the actor has few prior, ingrained attitudes that are inconsistent with the counter-attitudinal message, say, when the actor is a novice or an inductee in a new social environment, or when the perceived threat from counter-attitudinal groups is low;

- when the agent is relatively autonomous from the principal (e.g., when the issue is technical or ignored by the principal);

- when the actor is exposed to counter-attitudinal information repeatedly over time.

In practice, as I will come to in a moment, these conditions are more likely to hold in some kinds of institutions than in other kinds.[8]


### The ARF as a Counter-Realpolitik Institution

The first thing to determine is whether or not the ASEAN Way as embodied in the ARF does indeed constitute a counter-realpolitik ideology that is, in some sense, diffusible. Acharya identifies at least four key elements of this ideology: open regionalism, soft regionalism, flexible consensus, and cooperative/common security (Acharya 1997b).[9] The first three refer to the "structure" and form of the ARF, a variable that matters when discussing whether the ARF creates conditions conducive to persuasion (and perhaps social influence). I will come back to these features in a moment. Cooperative security, however, is the normative core of the ARF. First enunciated by the Palme Commission for Europe in the early 1980s, the concept embodies a number of principles: the non-legitimacy of military force for resolving disputes, security through reassurance rather than unilateral military superiority, non-provocative defense, and transparency. Behavior that is reassuring rather than threatening should be the rule, such that the ARF can "develop a more predictable and constructive

---

[8] Kelley (2004) attempts a head-to-head test of "rationalist" (e.g., material conditionality) versus "socialization" arguments (persuasion and social influence), but she does not really examine in any detail whether the conditions under which persuasion and/or social influence are most likely to pertain were present in the cases she examines. Indeed, her findings that the presence of powerful domestic opposition undermines international persuasion attempts does not undermine the argument about scope conditions for persuasion outlined here. One would expect that when domestic opposition forces are powerful, then the agent's autonomy should be relatively low, thus the agent should be less susceptible to persuasion attempts.

[9] Some analysts differentiate between cooperative and common security, but the differences are relatively minor and have to do with the issues that are considered security threatening (cooperative security uses a looser definition of security issues to embrace so-called nontraditional security—environment, social unrest, etc.). See Dewitt (1994).

pattern of relations for the Asia-Pacific region."[10] The security philosophy here implicitly assumes states are essentially status quo (or can be socialized to accept the status quo), and as such it is both normatively and empirically "true" that reassurance behavior is a better route to security than traditional realpolitik strategies. Security is positive sum. As such, traditional axioms like "If you want peace, prepare for war" are outmoded or counterproductive.[11] To this end, the normatively appropriate and empirically effective means for achieving security involve the building of trust through confidence-building measures, and the defusing of security problems through preventive diplomacy and conflict management. This is not to say that all members of the ARF, even the strongest backers of the institution, behave in ways perfectly consistent with the injunctive norms. The point is that these are the articulated and formal, if sometimes implicit, "theories" of security that are supposed to serve as the basis of "habits of cooperation."[12]

For a social environment to have a socializing effect, obviously an actor has to be a participant. The ARF is explicitly designed to be maximally attractive to states. The principles of open regionalism, soft regionalism, and flexible consensus are critical in this regard. Together they reflect the nondiscriminatory goals of the ARF. Although there are evolving rules for participation, the principle of open regionalism means the institution should be as inclusive as possible, combining multilateralist activists and skeptics such that there is no aggrieved actor left out to undermine the efficacy or legitimacy of the institution.[13]

---

[10] Most of these principles are embodied in the ARF Concept Paper (1995b). See also the comments by the Malaysian defense minister Hajib Tun Rajak cited in Dewitt (1994:12–13) and Lee Kwan-yew's comments about the ARF as a channel for China reassuring Southeast Asia about its status quo intentions in Makabenta (1994).

[11] "*Si pacem, parabellum*" in the Latin. "*Ju an si wei, you bei wu huan*" in the Chinese. These are the security principles of the OSCE as well. The primary difference between OSCE and ARF definitions of security is that the former includes human rights and liberal domestic governance as a component of interstate security. The ARF, sensitive to the postcolonial sovereign-centric ideologies in ASEAN and China, excludes this element.

[12] I do not mean to imply that the ARF was set up to undermine the balance of power ideology and practice, nor that many of its members are not consummate practitioners of realpolitik. My point is that the ARF ideology was not designed to promote balance of power ideology or practice. There are a range of views among ARF actors as to the complementarity between realpolitik and cooperative security. But the logic of the ARF ideology is to restrain the worst features of realpolitik, such as security dilemmas born from uncertainty or conflicts born from revisionist preferences. I thank Ralf Emmers for pressing me on this point. For a thorough discussion of the ARF's ideology, see Katsumata (2006).

[13] In this respect, the ARF reflects what Downs, Rocke, and Barsoom call a transformational regime, precisely the type of design that, they argue, is least conducive to effective multilateral constraints on behavior because it seeks out a lowest common denominator and dilutes the influence of "activists." Their argument holds *if* one assumes that preferences are fixed, that socialization does not occur, and that the ideology of the institution

Moreover, the institution should be as attractive to states as possible (in this case, China). Soft regionalism, therefore, emphasizes the informality and nonintrusiveness of the institution, and explicitly endorses the codes of conduct in the ASEAN Treaty of Amity and Cooperation (TAC), which emphasizes sovereignty-preserving principles such as noninterference in the internal affairs of states, respect for territorial integrity, the right to choose domestic social systems, and so on.[14]

At first glance this would appear to be inconsistent with counter-realpolitik socialization. I do not think there is any easy way of squaring this circle. What this principle does do, however, is send reassurance signals to participants that the institution will not undermine basic interests, that it will not be used by powerful states to exploit less powerful or influential ones. That is, it makes the institution attractive, or at least nonthreatening from the perspective of the most skeptical potential participant.[15]

Flexible consensus ensures both that the institution does not move far ahead of the interests of the most skeptical state *and* that the most skeptical state cannot easily veto its evolution.[16] Consensus decision making is a logical mechanism for reassuring member states that the institution will not threaten sovereignty or national unity. The rule was expressly written into the Chair's Statement summarizing the consensus at the second ARF meeting in Brunei in 1995: "Decision of the ARF shall be made through consensus after careful and extensive consultations among all participants" (ARF 1995a).[17]

---

is also diluted as the membership includes more "skeptics." It is not clear why this should be so, however, if the ideology is relatively stable and legitimate. See Downs, Rocke, and Barsoom 1997.

[14] On the TAC, see Leifer 1996:12–15. As I noted, these stand in contrast to the OSCE definition of common security. It is unclear how long the ASEAN Way discourse about security can resist a turn to the domestic sources of regional insecurity, however. The notion of nontraditional security issues—drugs, crime, refugees, transboundary pollution, and so on—has begun to enter the vocabulary of security specialists there. A number of ASEAN states are unhappy about how Burma's domestic governance performance negatively affects ASEAN's reputation. For a discussion of ASEAN's limited efforts to loosen the norms of noninterference, see Acharya 2004.

[15] These were norms that the Chinese regime, faced in particular with perceived American threats to unity (support for Taiwan) and domestic political order (human rights), wholly endorsed. As one Canadian diplomat noted, the ASEAN Way is a catchphrase for a pace that the PRC is comfortable with. The promise of a slow pace in the ARF is the only reason China came to the table (interview with Canadian embassy officials, Beijing, China, April 1996).

[16] The concept of flexible consensus came from ASEAN practice. See Acharya 2001:69.

[17] This statement, in turn, reflected the ARF Concept Paper, a blueprint for the ARF's institutional and agenda evolution. "The rules of procedure of ARF papers shall be based on prevailing ASEAN norms and practices: Decisions should be made by consensus after careful and extensive consultations. No voting will take place" (ARF 1995b:6).

Consensus decision making might appear to be a suboptimal decision-making rule for a diverse group of actors: although it is more efficient than a unanimity rule, there is always the risk that individual actors can acquire informal veto power.[18] Studies of consensus decision making among political parties in Swiss canton governments suggest, however, that consensus rules are likely to reduce intergroup conflicts in systems with "strong subcultural segmentation" (e.g., diverse subgroups as in the ARF) (Steiner 1974). In addition, as Chigas, McClintock, and Kamp argue in their analysis of consensus rules in the OSCE, consensus means all states have a greater stake in the implementation of decisions because they are collectively identified with a decision in ways that they would not be had they been defeated in an on-the-record vote over a particular course of action. Efforts to buck or shirk consensus decisions will generate more negative "peer pressure" than had clear opposition been registered through a vote (Chigas, McClintock, and Kamp 1996:42–43). Put differently, consensus rules make obstinacy costly in ways that up-and-down voting rules do not: abstinence threatens to undermine the effectiveness of the entire institution because its effectiveness is premised on consensus. It portrays the obstinate actor as one whose behavior is fundamentally at odds with the purposes of the institution. "Principled stands" against efforts to declare consensus are viewed as less principled than had they been expressed in a losing vote. Moreover, a consensus decision reduces the risk of ending up on the losing side. Losing internationally can have domestic political costs. It could be harder to maintain a domestic consensus for an international institution if one appears to lose badly from time to time (Steiner 1974:269–71; Lindell 1988:45).

The ARF's consensus decision rule was an attractive feature for China, even though as the ARF evolved it was hard to use flexible consensus norms for outright veto purposes. Consensus ensured that China would not be on the losing side in any majoritarian voting system. This was probably important for those in the Ministry of Foreign Affairs handling ARF diplomacy: it would have been much harder to sell the benefits of the ARF in the policy process in Beijing if China's leaders had evidence that China was losing in recorded voting procedures.

A subcomponent of consensus decision-making rules in the ARF is a norm of avoiding particularly controversial issues that might end up pre-

---

[18] The application of consensus rules in international organization varies a great deal. It can take the form of anything from a norm of unanimity in which there is informal vote taking and where one state can veto decisions, to a norm where the chair has such legitimacy and latitude that individual opponents to a declaration of consensus are reluctant to challenge. The ARF tends to operate more closely to the latter than the former. On variations in consensus practices, see Lindell 1988.

venting consensus. This is where Track II activities have been instrumental to the functioning of the ARF, both as a source of ideas and as a channel for defusing potentially volatile issues. These Track II activities come in three forms: ARF-sponsored Track II meetings;[19] activities undertaken parallel to, or in support of, the ARF without the ARF's prior formal endorsement;[20] and the Council on Security Cooperation in the Asia Pacific (CSCAP), an umbrella organization created in 1993 of thirteen national CSCAP committees. Although it is not the only Track II process around, CSCAP is the largest and most organized, with national CSCAP committees collaborating in working groups on topics such as CBMs.[21]

Whether or not by design, the evolving relationship to Track II contributes to the ARF's stability and legitimacy as an institution for states in the region. Issues that are too controversial for Track I can be moved into Track II rather than being discarded entirely. This sustains the momentum behind issues that the ARF might otherwise be compelled to abandon at the Track I level. Given that many Track II participants are government officials who also participate in Track I activities,[22] an issue is never really

---

[19] For example, in a Paris workshop in November 1996 on preventive diplomacy, the chair's statement recommended that the ARF consider taking a more proactive role in preventive diplomacy through the provision of the ARF chair's "good offices."

[20] For example, Australia convened a workshop on CBMs in November 1994; Canada and Malaysia cohosted a workshop on PKO activities in March 1995; and South Korea hosted a workshop on preventive diplomacy in May 1995. The results of the workshops were acknowledged and commended in the chair's statement at the 1995 ARF. See ARF 1995a:5 and Smith 1997.

[21] The relationship between CSCAP and the ARF has been rather ambiguous. Neither the 1995 nor 1996 ARF chair's statement specifically names CSCAP as the primary forum for ARF Track II activities, although the 1995 Concept Paper does identify it, along with ISIS ASEAN as two potential braintrusts for the ARF. Its absence from the chair's statements reflected, most likely, Chinese objections at that time to handing Track II responsibilities to an organization in which China was not a member. China's membership had been held up as the rest of CSCAP debated how to handle Taiwan's application for membership. The PRC refused to set up a national committee until it was satisfied Taiwan could not participate formally. This decision was made in late 1996; the PRC subsequently put together its national committee and formally applied to join CSCAP. CSCAP emerged as a potential ideas factory for the ARF, somewhat analogous to the non-governmental Pacific Economic Cooperation Council's relationship to APEC (interview 1997 and e-mail correspondence with an Australian government official involved in ARF policy making, February 1997). On the work of CSCAP in providing new ideas and proposals to individual ARF participants and to ARF working groups, see Simon (2001).

[22] For example, in the 1990s about 50 percent of the board of directors of the US national committee of CSCAP had worked in government. The US CSCAP also has a category called observers who are current government officials (US CSCAP 1997). The original Chinese CSCAP national committee was very small and exclusively reserved for senior officials. It included an assistant foreign minister, the senior specialist on American, European, and arms control affairs in the PLA General Staff Department, as well as the Foreign Ministry's senior functional-level officer handling ARF affairs (PRC CSCAP n.d. 1997?). In 2000 the

not within Track I's sphere of attention. This means that states can get more used to an issue being part of their interaction than if it were initially considered illegitimate. Track II can also "filter" or sanitize proposals that would otherwise be deemed more controversial by dint of who made them.[23] *Who* makes a proposal can sometimes be more controversial than the *content* of the proposal itself (Desjardin 1996). But if proposals are depersonalized through the Track II consensus process, and then again through the ARF chair's determination of consensus in the Track I level, much of the controversy can be filtered out. Thus, Track II can help define a Track I agenda that might not have otherwise appeared. As long as this myth of difference is not explicitly challenged, then the destabilizing effect of controversial issues is reduced. Chinese officials have stated openly that CSCAP's unofficial nature was a fiction because of the presence of so many government officials in their "personal capacities." Nonetheless, the Chinese government has played along: in a statement of support for links to Track II, it noted, "Issues not discussed or needing further discussions because of disagreement" can be put into Track II fora.

The form of the ARF, then, exhibits some of the features of an institution that may be likely to create a social environment conducive to persuasion: membership is relatively small (twenty-two states as of this writing) with some consistency over time in the participants at both the senior minister and functional specialists levels; the decision rule is (flexible) consensus; the mandate is deliberative and, partly as a result, this lowers the perception that highly threatening states can control the outcomes of the institution; and initially at least there was a certain amount of autonomy for China's representatives to the extent that the ARF was not central to Beijing's regional diplomacy, and the most likely repository of opponents, the PLA, was not fully involved in policy making.[24]

---

membership was expanded to include younger scholars and think-tank analysts who worked on regional security issues.

[23] "Filter" is Paul Evans's term. I am indebted to him for his insights into Track II. For an insightful discussion of the social psychological theory behind Track II effectiveness, with application to the Middle East, see Kaye 2001. For a discussion of the role of Track II in ASEAN politics, see Acharya 2001:66–67.

[24] The franchise characteristics of the ARF are hard to code. On the one hand, there is no formal recognition of particularly authoritative voices, e.g., there are no scientists' working groups or advisory panels that often define the scientific boundaries of policy discourse in, say, environmental institutions. On the other hand, ASEAN states are clearly authorized to take leadership roles in all ARF activities. All ARF intersessional meetings, for example, must be cohosted by an ASEAN state. The so-called ASEAN Way, therefore, is enshrined as the guiding ideology of the institution. One complicating factor, however, is that often ASEAN states can be quite passive in the promotion of the common security elements of the ASEAN Way, particularly when these conflict with its sovereign-centric elements. Thus, on transparency issues or intrusive confidence-building measures, China is not always the only state pushing for a lowest common denominator solution. On certain multilateral is-

We have, then, two key features of the ARF: a counter-realpolitik ideology and an institutional structure with features conducive to the development of "habits of cooperation" in the absence of material threats and punishments. On top of this, the institution is seen explicitly by many of its participants as a tool for socializing China to accept multilateralism, transparency, and reassurance as a basis for security.[25]

Put differently, some participants in the regional security discourse see the ARF as a tool for increasing China's "comfort level" with multilateralism. Comfort level is another way of saying that an actor's utility level changes positively with changing levels of institutionalization. An actor has a particular distribution of utility associated with particular levels of institutionalization.[26] Different actors may have different distributions of utility. Skeptics of multilateralism would have low values of utility at high levels of institutionalization. Committed multilateral activists would have high values of utility at high levels of institutionalization. Greater willingness to accept institutionalization would be indicated by an increase in an actor's utility whereby it comes to believe that the absence of an institution is less valued than before and the presence of one becomes more valued than before.

The question is, what might cause a shift in this comfort level in this distribution of utility? Mainstream institutionalist theory would probably focus on things such as reassurance (information that underscores fears of even small amounts of institutionalization are exaggerated) or the distributional effects of the institution (leading to change in domestic political balances of power). Socialization arguments would focus on persuasive arguments that more institutionalization is a "good" in and of itself, or on social backpatting and opprobrium effects that link the utility of involvement in the institution to the utility of social status and diffuse image.[27]

---

sues, non-ASEAN activist states take the lead in defining the discourse, e.g., given its experience, in intersessionals on peacekeeping operations Canada has spoken with a more authoritative voice.

[25] I do not want to leave the impression that everyone in ASEAN intended to try to alter Chinese interests. Some were more skeptical about this possibility than others. Even these people, however, did not necessarily have an interest in openly defining the security problem as something akin to a suasion game. The concern was that by focusing on a China threat, the ARF will lose its status as a focal point and ASEAN will lose its leadership status in regional security affairs.

[26] I am subsuming, for the moment, the intrusiveness of the agenda within the level of institutionalization.

[27] I would call this shift in utility distribution a change in preference. I realize some contractual institutionalists would debate this, and consider this a change in strategy. The difference between the two concepts is artificial and depends on the level of ends and means one is examining. For game theorists, the outcome of strategic interaction between two players is the product of a particular strategy pair. States are said to have preferences over outcomes.

Here I want to focus on evidence for persuasion. Recall the required indicators of persuasion: that social environments in the institution are conducive to persuasion; that after exposure to or involvement in a new social environment, attitudes or arguments for participation converge with the normative/causal arguments that predominate in the social environment; that behavior has changed in ways consistent with prior attitudinal change; and that material side payments or threats are not present, nor are they part of the decision to conform to pro-social norms.

Having established that the institutional form of the ARF meets the criteria for an environment conducive to persuasion, the question is whether, then, attitudes or arguments in China converged with the normative/causal arguments at the core of the ARF "ideology." It is quite clear that the public and internal discourse in China on multilateral security dialogues in the Asia Pacific prior to China's entry into the ARF in 1994 was highly skeptical of their value. Indeed, in internal circulation (*neibu*) and open materials alike, the discourse stressed that bilateral relations, particularly among the great powers, were the basis of stability or instability in IR; that there was no urgent need to build multilateral security mechanisms, indeed that multilateralism was unimportant for handling regional problems; that such institutions would be dominated by the United States or Japan while China would be outnumbered; and that sensitive bilateral disputes where China might have an advantage in bargaining power might be internationalized.[28] The skepticism of multilateralism was rooted in even deeper realpolitik assumptions about international relations where structurally (and sometimes ideologically) induced zero-sum competition among sovereign states necessitates unilateral security strategies.[29]

From the mid-1990s on, however, there were some noticeable changes in the discourse. For one thing, China's exposure to regional multilateral institutions led to a dramatic intensification of the discourse about regionalism and multilateralism. As figures 4.1 and 4.2 show, there has been an

---

Yet if an institution is itself a product of a particular strategy pair (or the confluence of more than two strategies in a multilateral institution), then the form and function of the institution itself is a preference. Of course, multilateralism can also be a strategy at a higher level of interaction where the goal is some more abstract good, such as security. But since security is a grand preference of most states, to limit preferences to things as abstract as security, welfare, peace, and so on, means that no outcome below this level can be called a preference. Everything becomes strategy. I think this reduces the utility of the term "preference," and ignores the fact that actors can come to internalize multilateralism, unilateralism, or bilateralism as legitimate, taken-for-granted ends in themselves.

[28] Interview with Chinese intelligence analyst involved in the ARF policy process, Beijing, July 1996; Garrett and Glaser 1994; Yuan 1996; Johnston 1990; Xu 1996:252–53.

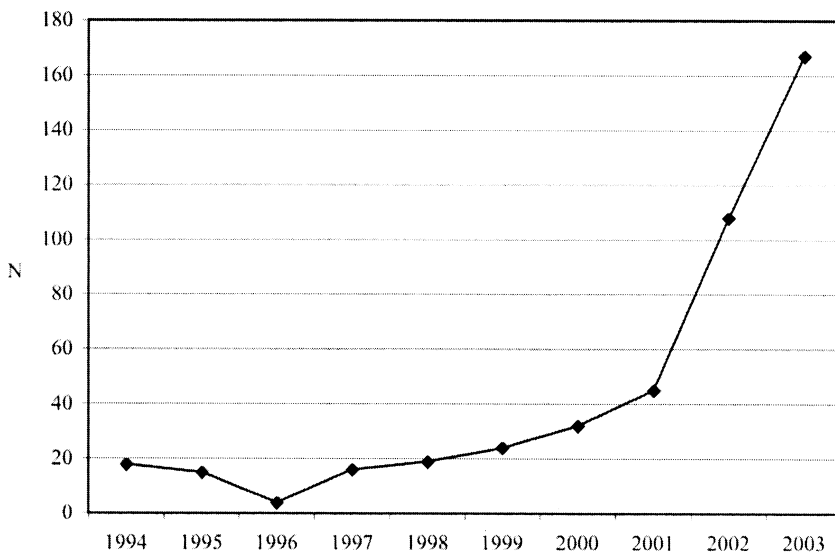[29] On Chinese realpolitik, see Christensen 1996 and Johnston 1998a.

Figure 4.1. Frequency of articles with the term "multilateralism" (*duobianzhuyi*) in Chinese journals on international relations. Source: China Academic Journals database, China National Knowledge Infrastructure.

exponential increase in the discourse about "multilateralism" (*duobian-zhuyi*) in academic and official publications over the past decade. References to the ASEAN Regional Forum have also increased substantially since the organization came into being (figure 4.3). Assuming frequency has some face validity as an indicator of attention paid to an issue, it is clear that interest has increased over time.

In addition to changes in the frequency of the discourse about multilateralism, we can observe a change in its content as well. Initial statements made to the ARF (such as Foreign Minister Qian Qichen's comments at the first ARF in 1994) stressed what can only be seen as traditional "rules of the road" for the management of relations among sovereign, autonomous states. These included the Five Principles of Peaceful Coexistence, economic ties on the basis of equality and mutual benefit, the peaceful settlement of disputes, and adherence to the principle that military power should be used only for defensive purposes (Yuan 1996:11). Terms, concepts, and phrases associated with common or cooperative security were absent.[30]

[30] See, for instance, one of the earliest analyses of the ARF by a Chinese think-tank specialist in Asia-Pacific security, Liu 1996.
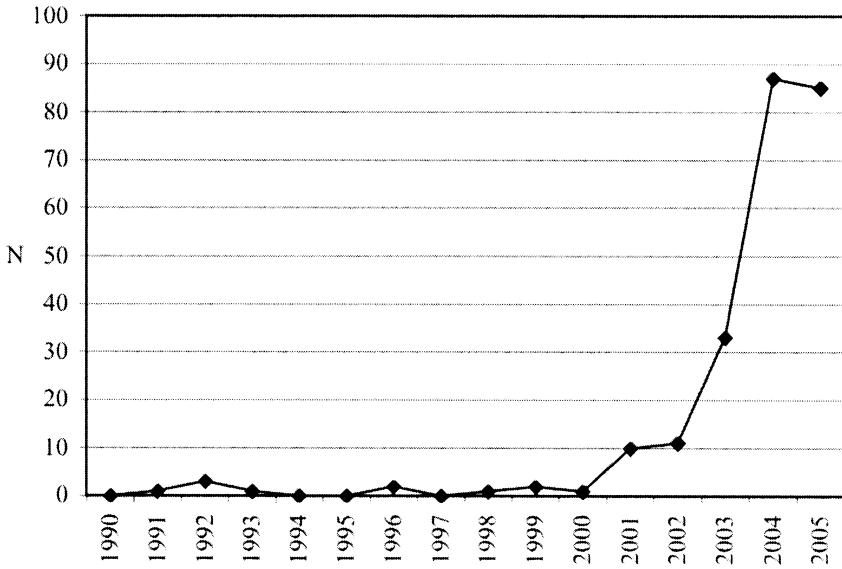
Figure 4.2. Frequency of articles with the term "multilateralism" (*duobianzhuyi*) in the *People's Daily*. Source: *People's Daily*, on-line full-text version.
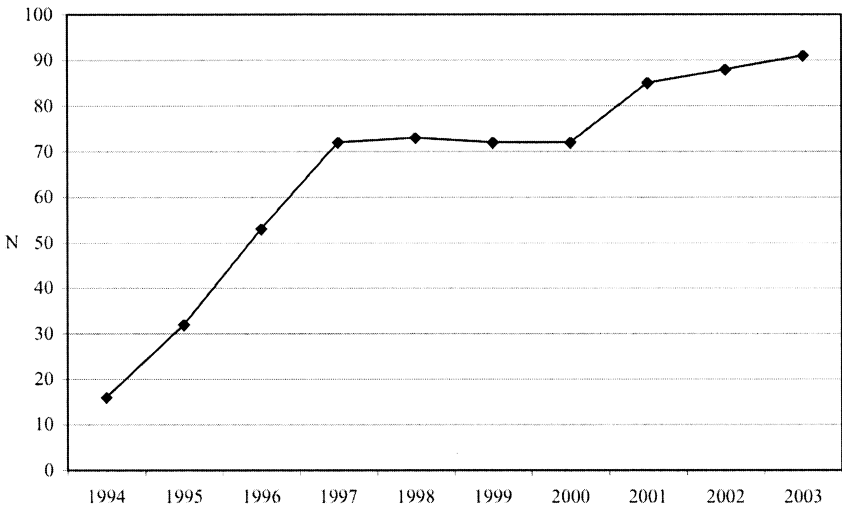


Figure 4.3. Frequency of articles in Chinese journals on international relations with references to the ASEAN Regional Forum (*Dongmeng diqu luntan*). Source: China Academic Journals database, China National Knowledge Infrastructure.

By late 1996, however, Chinese working-level officials directly involved with ARF-related affairs began to articulate concepts that were, to a degree, in tension with traditional realpolitik arguments. Shu Chunlai (a former ambassador to India, and a key figure on China's CSCAP committee) appears to have been China's first authoritative participant in ARF-related activities to have used the term "common security." In a paper originally presented at the ARF-sponsored Paris workshop on preventive diplomacy (November 1996), Shi and coauthor Xu Jian noted that common security was central to the post–cold war need for a "renewal" of old security concepts. This renewal, they argued, entailed abandoning "old" concepts, "based on the dangerous game of balance of power." There was not much more on the subject, and the paper went on to stress, somewhat in tension with common security, that preventive diplomacy should be handled strictly in accordance with the Five Principles of Peaceful Coexistence (Shi and Xu 1997).

By early 1997, however, ARF-involved analysts and officials *unofficially* floated a more well-developed concept of "mutual security" at the first Canada-China Multilateral Training Seminar (the seminar brought together a small number of key officials handling the ARF in the MOFA Asia Department, and a couple of analysts from China's civilian intelligence institution [CICIR] who were also in the ARF "interagency" process). The term meant, according to one Chinese participant, that "for you to be secure, your neighbor has to be secure," a common security concept based on the notion of "win-win." It is possible the Chinese may have felt under pressure to develop an original Chinese contribution to the multilateral security discourse: "common security" was perhaps too closely identified with the CSCE process, and thus might have been too provocative inside the Chinese policy process.

One of the participants in the seminar (a participant in interagency discussions on the ARF, and an analyst in CICIR) also submitted a paper for discussion in which he listed three types of security systems: hegemonic systems, alliance or military-bloc systems, and multilateral systems. The latter he called an "encouraging development" and noted that mutual security, like common security, cooperative security, and comprehensive security, were traditionally unfamiliar concepts in China. But these were now "taking place in the minds of policymakers and scholars and in the actions of Chinese policies," though he did not elaborate beyond this (Chu 1997).

Around the time of the seminar, another analyst involved in ARF-related work in a think tank attached to the State Council wrote a paper on confidence building in the Asia Pacific. The paper provided a sophisticated explanation of Western theories of CBMs, noting, for example, the military reassurance purposes of CBMs. The author also elaborated a bit on

"mutual security," noting that the concept was embodied in the April 1996 Five-Power (China, Russia, Kazakhstan, Kyrgyzstan, Tajikistan) Treaty on CBMs (Liu 1997). (One of the Chinese participants at the Canada-China seminar in Toronto had also noted that mutual security had come from the Chinese discourse on the Five-Power Treaty.)

The invocation of the Five-Power Treaty in this common security–influenced discourse on mutual security was important. The treaty comes as close to a CSCE-type CBM agreement as anything in the Asia-Pacific region, with provisions for limits on the size and type of military maneuvers allowed within certain distances of borders, provisions for military observers and military exercises, and so on.[31] In internal Chinese debates over multilateralism, whether or not one believed the principles of the treaty had broader applicability to the region was an indicator of one's skepticism toward multilateralism in general. Those who supported its applicability tended to be stronger supporters of multilateralism.[32] The initial idea for the treaty grew out of bilateral PRC-Soviet negotiations over the border in 1990 and 1991. The Soviets had introduced the idea of a formal CBM agreement, using conventional force reduction CBMs that they had negotiated with Western Europe as a template. Initially, Chinese negotiators were literally unsure of the meaning of the terms the Soviets were bringing over from the European experience. The terminology had to be translated into Chinese with explanation so that MOFA officials understood the implications of certain CBM terminology. Thus, the Five-Power Treaty emerged fairly directly from European CBM experiences tabled by the Soviets.[33] That those articulating the concept of mutual security would do so by invoking the Five-Power Treaty as an example/precedent/exemplar suggests that the term signified an acceptance of more intrusive and formal security institutions. Indeed, the earliest analyses of the ARF tended to explicitly reject the CSCE as a model for the Asia Pacific (Liu 1994:18).

[31] For a systematic comparison of the Five-Power CBM agreement and the CSCE Vienna document of 1994, see Acharya 1997a:16–23.

[32] This was the distinct impression I received when interviewing military and civilian specialists on the ARF. Classic examples were two senior military officers, one a longtime participant in arms control policy, the other in charge of ARF and regional security affairs in the PLA. Both believed that the principles of the Five-Power Treaty had at least some applicability to the rest of the region. This contrasted with other PLA officers with whom I spoke who were doubtful about applicability and about multilateralism in the region. I never met anyone who could be placed in the other two "cells" of this rough $2 \times 2$: namely, people who believed these principles were applicable and were hostile to multilateralism in the Asia Pacific, or people who believed these principles were not applicable and yet were strong supporters of regional multilateralism.

[33] Interview with Chinese arms control specialist, January 1999.

Interestingly enough, in June 1997 at the first CSCAP General Meeting, when China's national committee participated for the first time, China's representative, Chen Jian (assistant foreign minister, and formally in charge of multilateral security issues at the MOFA), explicitly extolled the Five-Power CBM Treaty as contributing to confidence and security in the region. He did not mention "mutual security," however (nor indeed the term "new security concept") (Chen Jian 1997). This suggested that there was still probably some internal debate about the legitimacy of the concept and whether China should be formally and publicly associated with it. The term had not yet made it into the official policy discourse.

This began to change in the second half of 1997. At an ARF Track II meeting on preventive diplomacy in Singapore in September 1997, Ambassador Shi Chunlai, one of the regular Chinese participants in regional security dialogues, suggested in response to a question that the CBMs China had developed with India and Russia were applicable to Southeast Asia.[34] Then in November 1997, the Chinese paper presented to the ARF Intersessional Support Group on CBMs in Brunei explicitly noted that the Five-Power Treaty embodied the notion of "mutual security" and could be used as a source of ideas for the rest of the Asia Pacific. Mutual security was defined as an environment in which the "security interests of one side should not undermine those of the other side. . . . This kind of security is a win-win rather than zero-sum game" (China 1997:3). We should not underestimate the significance of the incorporation of this loosely game-theoretic terminology into the Chinese discourse (and another, "positive sum"—*zheng he*—used more recently by multilateralists in China)—terms borrowed, one assumes, from interactions with multilateralists in ARF-related activities. The origins are hard to pinpoint, but the term "win-win" does not seem to have been used much prior to 1997 in discussions of regional security; one of its earliest appearances was in comments that some Chinese participants made in the Canada-China seminar in January 1997.[35] The term "win-win," of course, stands in distinct tension with traditional realpolitik notions of security and reflects core assumptions of common security.

Then in December 1997 at the Third CSCAP North Pacific Meeting, Ambassador Shi Chunlai developed the "new security concept" further, linking it to "mutual security" and, by implication, to common security:

[34] Notes on Shi Chunlai's presentation, ARF Track II meeting on preventive diplomacy, September 9–11, 1997, Library of the Institute of Defence and Strategic Studies, Singapore.

[35] The term was first used in a *People's Daily* article in 1993 in reference to China-Taiwan talks in Hong Kong. Its usage, and its spread into other discursive spheres, took off in 1998–1999. The earliest academic article to use the term in reference to East Asia security appears to be Ruan Zongze's analysis of Clinton's East Asia strategy. See Ruan 1996.

the concept, he argued, was "one that is not based on the cold war mentality featuring zero-sum game, but on mutual and equal security." Rather it meant "not creating winners and losers" (Shi 1997:1–2). Both "the new security concept" and its component "mutual security" received the highest-level endorsement when they were included in remarks by China's foreign minister Qian Qichen at the Private Sector's Salute to ASEAN's Thirtieth Anniversary in December 1997. Since then the concept has been incorporated into China's ARF discourse. In the words of one senior diplomat engaged at that time in the ARF diplomacy, the concept comes close to cooperative and common security but stresses trust and confidence building more than institution building.

Over time, official Chinese commentary pushed the discourse further to include rather direct attacks on realpolitik concepts. An analysis broadcast by China Radio International in late December 1998 argued, for instance, that the Five-Power CBM Treaty was a good example for the rest of the Asia Pacific. It had authenticated "a new security concept completely different from the cold war mentality and the traditional security concept, if you desire peace, you must prepare for war. This saying is a vivid description of the traditional security concept." The traditional realpolitik concepts included ideas such as maximizing military force so as to become stronger than one's opponent, a narrow focus on the security of the nation above all else, and the resort to military means in the pursuit of security (China Radio International 1997).[36]

Such talk could be dismissed as rhetorical bandwagoning by state-controlled media. But it reflected an incorporation of language that at least some of its users saw as challenging traditional state practices. More significant, however, was the interest in MOFA's Asia Department in multilateralism. Realizing that it required more sophisticated arguments to bolster and justify the mutual security discourse and policy, the MOFA began to ask some of the key IR thinkers in China, a number of whom one might consider multilateralists and integrationists, how to put more theoretical

---

[36] The explicit rejection of the parabellum phrase is, in a sense, also a repudiation of a long-standing Chinese equivalent: *Ju an si wei, you bei wu huan* ("When residing in peace, think about danger. With [military] preparations there will be no calamities"). This does not mean that mutual security is a fully developed concept, nor that it wholly replicates cooperative or common security concepts. According to interviewees in October 1998, the terms "mutual security" and "new security concept," although developed by the Asia Department of the Ministry of Foreign Affairs precisely in response to the requirements of participating in multilateral activities in the region, were not fully developed themes in Chinese foreign policy. In 1998 the MOFA tasked specialists in the Chinese Academy of Social Sciences to develop further the theoretical cores of these concepts. Interviews with Chinese think tank analysts involved in the development of the new security concept, October and November 1998; conversation with MOFA official, October 1998.

meat on the conceptual bones.[37] To this end, the MOFA's Policy Research Office, on behalf of the Asia Department, commissioned a study by a respected specialist in regional multilateralism from the Chinese Academy of Social Sciences. The report was entitled "The Concept of Comprehensive Security and Some Theoretical Thoughts about China and Asia-Pacific Security" and was submitted to the MOFA in December 1998. The author, Zhang Yunling, explicitly argued that military power and traditional territorial-based concepts of national security were no longer the most important issues in China's future security in the region. Rather, China faced an increasing array of nontraditional security problems that could not be solved through the augmentation of national military power alone, and thus should focus more energy on developing multilateral cooperative solutions to security problems, including greater activism in the ARF. The report noted—in recognition of security dilemma dynamics—that China's behavior on the ground was one reason why states worried about China's rising power. To deal with this, the report argued, China had to signal that it basically accepted extant rules of international and regional order, while trying to moderate these rules and norms through existing international institutions and procedures. In other words, China's rise was a potentially destabilizing element in international relations because of perceptions of Chinese power in the past, but that China had to credibly signal that it was in essence a status quo power. The report explicitly borrowed arguments and concepts from Western, including Canadian, multilateralists and included an appendix that introduced some of the multilateralist lexicon to its audience (e.g., integration theory, interdepen-

[37] Interview with senior MOFA official engaged in ARF work, January 12, 1999. These individuals included Yan Xuetong, Wang Yizhou, and Zhang Yunling, among other well-respected IR specialists. Yan is considered somewhat more realpolitik in his views of multilateralism than Wang or Zhang. Nonetheless, he has argued that the new security concept and mutual security are based on common security and constitute a recognition that security dilemmas are a major source of conflict in the international system (Yan 2005:163). Zhang, an economist, views regional security through the lens of economic integration. Wang is one of China's most influential IR specialists, with a research interest in the process of China's integration into global institutions. His work has been influenced by his exposure to liberal institutionalism and social constructivism. The effort to develop more sophisticated thinking about multilateralism was given a boost in 2001 with China's first conference on the topic, hosted by the Institute of World Economics and Politics at the Chinese Academy of Social Sciences. Admitting that the topic was still very sensitive in China, one of the organizers, Wang Yizhou, wrote that multilateralist theory requires China to rethink its opposition to participation in everything from the G-8 to formal relations with NATO, to participation in ASEAN-US multilateral military exercises. See Wang 2001. Within a short few years, China in fact participated as an observer in the G-8, and broached the question of more formal relations with NATO.

dence theory, democratic peace theory) (Zhang 1998). The arguments in the report were designed in part to assist the Asia Department and other multilateralists to make more sophisticated internal arguments in favor of greater participation in the ARF, that is, to persuade others in the policy process, particularly in the PLA, of the value of multilateral diplomacy.[38]

The discourse on multilateralism in China, then, moved quite some distance over the 1990s, from public skepticism to informal articulations of mutual security and common security to public affirmation of the concepts. The concepts were explicitly linked to real-world institutional exemplars of these principles, such as the Five-Power CBM Treaty. This document was consistent with, indeed modeled in some ways from, CSCE-style institutions. For the protomultilateralists, the Five-Power Treaty was an example of applied mutual security.[39]

Moreover, it appears that those most directly participating in these various regional multilateral security institutions tended to express a higher level of valuation for them than those who had at best an indirect involvement or no involvement at all. As an initial test of this possibility, I analyzed a selection of ARF- and CSCAP-related documents, statements, position papers, and working papers produced by Chinese participants, as well as a collection of articles on regional multilateral security authored mostly by people who are not multilateral security specialists per se nor especially active in these institutions. The latter collection came from two policy journals in Beijing—one published by a Chinese military intelligence unit, the China Institute of International and Strategic Studies (*International Strategic Studies*), and one published by a civilian intelligence unit, the China Institute of Contemporary International Relations (*Contemporary International Relations*). Using a computer-aided content analysis program, I compared the former set of documents with the latter, looking in particular for differences in the valence attached to "multilateralism" and related terms in the texts. The measure of valence was the number of positive words as a percentage of the total number of positive and negative words semantically associated with terms for multilateralism. As it turns out, references to multilateralism were indeed more posi-

[38] Interview with Chinese academic specialist on Asia-Pacific multilateralism, January 1999. The report was published about a year later in *Dangdai Ya Tai* (Contemporary Asia-Pacific Studies). See Zhang 2000. Interestingly, even in the post-Kosovo atmosphere in China where more "liberal" voices had to tread with somewhat more caution when discussing international relations, there were essentially no significant changes in the wording or argumentation in the published version of the study.

[39] Interview with senior MOFA diplomat involved in ARF work, January 1999. The diplomat acknowledged that the treaty had been influenced by the OSCE model. The diplomat also acknowledged that the OSCE model was at least partly applicable to the Asia Pacific.

tive in the materials written by active participants in regional security institutions compared with those who were generally outside analysts.[40]

The obvious question is, has all this changing discourse about multilateralism been just cheap talk? A realpolitik actor would have incentives to use such a discourse deceptively: if one believed one was in a prisoners' dilemma environment, then cooperative discourses could encourage others to cooperative, thus creating opportunities to acquire the "temptation" (C,D) payoff. This would, in principle, be especially attractive to an actor in an institution such as the ARF, with little or no monitoring capacity (except for voluntary and nonstandardized "defense white papers") and no ability to punish defection. Some in the US government viewed, and continued to view, the multilateralist and mutual security discourses precisely as that: a deceptive effort to redirect attention from inconsistencies between Chinese security behavior (such as sharp increases in military expenditures and provocative military exercises) and the ideology of genuine multilateralism, while trying to underscore the incompatibility of US bilateral alliance strategies with multilateralism in the region.[41]

---

[40] The materials from those most active in multilateral institutions were collected in 2003 while I was at the Institute of Defence and Strategic Studies in Singapore. The IDSS has relatively complete records of both ARF and CSCAP meetings starting in 1994. I cannot guarantee I collected the universe of texts, but I believe the number I analyzed ($N = 33$) is fairly representative. The articles from the two think-tank journals were collected according to titles that mentioned Asia-Pacific multilateral security or related topics beginning in 1994. The $N$ for this group was smaller (13). Both sets of documents were entered into the Yoshikoder, a computer-aided text analysis program that allows one to isolate specific terms in the text and then look for positive or negative valenced words within a certain distance of the term of interest. The intuition that generally consistent characterizations and descriptions of a word can be found within a limited distance from the word comes from semantic space analysis (Lowe 2000). In this case, I looked for positive and negative terms within ten words of the term "multilateralism" and related words such as ARF, CSBM, ASEAN, CSCAP, preventive diplomacy, multilateral mechanism, and cooperative security. The dictionaries of positive and negative terms were drawn from the General Inquirer dictionaries. For the ARF- and CSCAP-related texts, the number of positive terms as a portion of the sum of positive and negative terms was 0.91, whereas for the journal articles it was 0.86. A one-tailed t-test shows that the difference is statistically significant at the 0.1 level ($p = 0.06$). I am very grateful to Will Lowe for designing and writing the Yoshikoder program.

[41] The argument is outlined in Garrett and Glaser 1997. See also Finkelstein's contention (1999:43) that much of the Chinese discourse on multilateralism was designed to counter US security bilateralism around the world and particularly in Asia. My own conversations with Pentagon officials involved in Asia policy confirms this particular interpretation. What lent this argument credence was that the mutual security discourse emerged around the same time that the Chinese hosted a rather uncharacteristically contentious (for the ARF) intersessional support group meeting on CBMs in March 1997. The Chinese diplomat chairing the meeting (actually cochairing with the Philippines, though the Philippines played a passive role in the meeting) refused to drop a Chinese agenda item that called for the study of CBMs (military observers, prior notification, etc.) at joint military exercises in the region. Since the United States and its allies conduct joint exercises while the Chinese do not, the

I am not convinced of the pure deceptiveness of this discourse, however. In principle, there is a relatively easy test of this hypothesis. If it were right, then the strongest proponents of the mutual security discourse (and multilateralism in general) and the Five-Power Treaty as an exemplar agreement for the region should have been the strongest opponents of US bilateral alliances in the region. In addition, variations in Chinese efforts to undermine support for US alliances (particularly with Japan) should have tracked directly with variations in the strength and prominence of the mutual security discourse.

On both of these tests, the instrumental or deception hypothesis comes up short. A careful tracking of the discourse, as I have tried to do earlier, suggests that originally the strongest proponents were precisely those who in their private interactions within the Chinese policy process, and with foreign diplomats and observers, indicated a deeper commitment to multilateralism. The proponents included multilateral functional specialists in the MOFA and moderate voices and "new thinkers" in the strategic analysis community. Although these people were generally opposed to the expansion of US-Japanese security cooperation, and wanted to use multilateral diplomacy to pressure the United States to limit the scope of its military cooperation with Japan, they also recognized the alliance was a reality and could indeed constrain Japanese remilitarization. Indeed, Zhang Yunling's MOFA-commissioned report on comprehensive security, mentioned earlier, was explicit in stating that China should not and need not replace US military superiority in the region, and that China need not balance militarily against US power, especially if effective, practical multilateral security institutions could be set up in the region.[42]

The proto-multilateralists have also argued that China's integration in multilateral institutions can help reassure the United States (Tang and Zhang 2004:6).

---

proposal was rightly criticized as being aimed at US military interests. The agenda item, however, was drafted by the PLA. The chief MOFA ARF policy functionary had privately indicated a willingness to drop the issue in discussions with the Canadians a couple of months prior to the meeting. But it is plausible that with the meeting in Beijing, and with a large PLA contingent observing the discussions, the MOFA did not feel free to drop the issue. In any event, the Chinese insistence on maintaining the agenda item in the face of opposition from a range of states prevented consensus on this issue and led to a great deal of concern in the United States about a Chinese offensive against US military alliances in the region. Conversations with Canadian diplomats, May 1997, and a PLA officer involved in the ARF policy process, December 1997.

[42] See Zhang 2000. See also Wang Yizhou's essay (2001) in which he advocates under the banner of multilateralism China's participation in ASEAN-US joint military exercises. See also Tang and Zhang 2004:5 in which they note that China should be able to coexist with the United States as a hegemon, as long as it does not threaten core Chinese interests. All of these authors have advised the MOFA's Asia Department on regional security issues.

In contrast, my sense is that those who were less enamored with the mutual security discourse were found mostly in the military, and it was in the PLA where some of the strongest skeptics of the US-Japan alliance were and are found. One could also imagine that the PLA should have been troubled by the anti-realpolitik content of "mutual security" and its use of a potentially militarily intrusive CBM treaty as a model for the region. Moreover, the Chinese CBM proposals in the ARF that were clearly biased against US military power in the region (e.g., observers at joint military exercises, reductions in military reconnaissance activities aimed at ARF members, etc.) appeared first in 1995–1996, well before the "mutual security" concept emerged, and were promoted by the PLA, not the Ministry of Foreign Affairs.[43] The disconnect is further underscored by the fact that the shrillness of the anti–US-Japan alliance rhetoric died down in the late 1990s, even as the mutual security concept was given higher profile and became more sophisticated in public discourse. This is inconsistent with the deception hypothesis about the origins of the multilateralist and mutual security discourses.

My argument here rests, obviously, on the critical question of whether the mutual security discourse has, to some degree, been internalized among those working most closely in the ARF environment. This appears to have been the case.[44] China's involvement in the ARF and related processes seems to have led to the emergence of a small group of policy makers and analysts with an emerging, if tension-ridden, normative commitment to multilateralism because it is good for Chinese and regional security. One longtime participant observer of multilateral security dialogues in the region, Amitav Acharya, described in a report to the Canadian government in the late 1990s how Chinese participants had developed a more positive evaluation of the role of multilateralism in improving Chinese security. Multilateral institutions such as the ARF helped Chinese leaders understand better how other states in the region perceived Chinese power (an understanding that is the first step in recognizing the potential for security dilemmas). According to one of Acharya's informants, China was "learning a new form of cooperation" (Acharya 1998:4). Another remarked, "Because of the ARF, China is more willing to settle its disputes by peaceful means" (Acharya 1998:5).

---

[43] These proposals came from the Comprehensive Department of the Foreign Affairs Bureau of the General Staff Department, the department that coordinates PLA positions on the ARF.

[44] Interviews with Canadian and Singaporean embassy officials, Beijing, April 1996 and October 1998; Chinese analyst involved in ARF policy, Beijing, July 1996; and Canadian specialist on Track II, January 1997; Smith 1997.

Another keen observer and organizer of some of these dialogues, Susan Shirk, was even more specific about how participation affected two of the leading MOFA supporters of multilateralism in China's Asia policy:

After participating in NEACD [Northeast Asia Cooperation Dialogue] and the official multilateral groups like the ASEAN Regional Forum that were just starting, Fu Ying became an advocate within the bureaucracy for regional multilateral cooperation. As she moved up the career fast track in the Asia Department, China's approach toward regional cooperation became increasingly pro-active. Fu succeeded in bringing Wang Yi (then head of the Asia Department) on board with multilateralism, as I learned when she arranged for me to talk with him after the 1996 NEACD meeting in Beijing. A dashing Japanese speaker, Wang Yi later visited me in San Diego, spent six months at Georgetown University to prepare for future promotions by improving his English, and enrolled as a part-time PhD student at the university attached to the foreign ministry. This bold and bureaucratically nimble duo brought a new spirit of confident cooperation to China's Asian diplomacy. (Shirk 2007:190)

ARF policy in China was initially put in the hands of the Comprehensive Division of the Asia Department of the Foreign Ministry. The division had only about eight to ten overworked officers. Since the Comprehensive Division normally handled a range of functional issues in addition to security issues, only a couple of these officers initially did the preparatory work for ARF meetings and Track II activities. In early ARF activities the Chinese representatives were unaccustomed to the give and take of corridor debate and negotiation. They also came to the discussions with a watchful eye for developments that might impinge on sensitive security or domestic political issues. Over time, however, with experience in informal discussion and familiarity with the ARF norms of interaction, these officers became much more engaged, relaxed, and flexible. Participants from other countries have remarked on the generally sophisticated, nuanced, amicable tone of Chinese interventions in discussions over time. Even Chinese ARF specialists have noted that the institutional culture of the ARF requires them to adjust the tone and tenor of their discourse. Unlike in the UN where vigorous and legalistic defenses of specific positions in negotiations that are often viewed as close to zero-sum are often required, in the ARF there is more give and take, more spontaneous intervention to explain positions, and with some exceptions, an atmosphere that downplays "in your face" defenses of national positions.[45]

---

[45] Interview with senior MOFA official involved in ARF diplomacy, January 12, 1999; interview with Canadian diplomats involved in ARF work, July 3, 2001.

It is important to note in this regard as well that in some instances, those exposed to multilateral security environments took on positions/ orientations that were distinguishable from others with their own bureaucratic background. It is clear, for instance, that the two PLA officers in the Foreign Affairs Bureau (later Office) who handled the PLA's input into regional security dialogue policy in the 1990s and early 2000s were more favorably disposed to multilateral solutions than the PLA as an institution.[46] Indeed, rare for a PLA officer, at least one of the two agreed that the principles of the Five-Power CBM Treaty of 1995 were applicable to East Asia more broadly.[47] As I noted, one's position on the wider applicability of this treaty was a reasonably good indicator of one's support for security multilateralism in the region.

Most interesting was this protomultilateralist constituency's apparent endorsement, within limits, of multilateralism as being compatible with Chinese security interests. Some foreign diplomats in Beijing who had interactions with these MOFA officers extensively suggested that their agenda was to tie China gradually into regional security institutions so that someday China's leaders would be bound by them. These officers saw involvement in the ARF as a process of educating their own government. Some even remarked that involvement in the ARF had reduced the likelihood of China's resort to force over disputes in the South China Sea because there were now more diplomatic (that is, multilateral) tools at China's disposal.[48] More generally, some Chinese analysts saw the ARF experience as contributing to a greater understanding of the notion of common security. As one article put it, in the context of the ARF's impact on regional politics, "following the growing intimacy of relations among countries in the Asia Pacific, and the increase in international responsibilities after the improvement in China's international position, China also became aware that its own security was not an isolated issue. At the same time as it strengthens its national defense modernization, it must rely on equal dialogues and discussions, and with other countries together discuss issues concerning the establishment of common security" (Liang and Zhao 2001:43).

The main conduit for the infusion of these sorts of ideas into the group of protomultilateralists was experience in Track I and II activities.[49]

---

[46] Interview with senior PLA officer in charge of regional multilateral security policy in the PLA, January 1999; and interview with Canadian diplomats involved in regional multilateralism, July 2001.

[47] Interview with senior PLA officer in charge of regional multilateral security policy in the PLA, January 1999.

[48] Interview with Canadian embassy officials, Beijing, April 1996; interview with Canadian embassy official, Beijing, October 1999; Acharya 1998.

[49] Interview with Chinese intelligence analyst involved in the ARF policy process, Beijing, July 1996.

It seems this group's influence over Chinese ARF policy was helped by further institutional change in China. In January 1998, for instance, the Asia Department set up a separate division just to handle ARF and Track II diplomacy. Interestingly, as one might expect, the creation of this special regional security institution inside the MOFA also led to an emergent organizational interest in ARF diplomacy. As one MOFA interviewee implied, as the ARF agenda moves toward considering more formal arms control like CBMs, the Asia Department has had to defend its prerogatives against the Arms Control and Disarmament Department, which handles most other multilateral security diplomacy for the MOFA.[50]

There is some intriguing evidence of the commitment these individuals had in protecting the policy from domestic political critics—hence, an indication of their growing normative stake in the ARF. A senior Canadian official involved in ARF diplomacy reported that the Chinese delegates to early ARF discussions apparently did not report back to Beijing references by other delegations to the CSCE/OSCE as a possible model for the ARF. The OSCE was not just a symbol of a more intrusive, constraining regime, it was also a regime that dealt with human rights (Smith 1997:18). Downplaying this information, then, was important for preserving support or acquiescence inside China for further institutionalization of the ARF. Canadian and American diplomats also reported that sometimes China's ARF delegates would help other states frame proposals for ARF-related activities in ways that would make these more acceptable in Beijing.[51] When the ARF diplomats were under closer scrutiny from Beijing, they tended to be less conciliatory publicly. During the 1997 ISG on CBMs in Beijing, for instance, Canadian and American diplomats observed that the MOFA diplomats stuck to the proposal for observers at joint military exercises due, possibly, to the large presence of PLA observers in the meetings. The MOFA ARF diplomats had earlier suggested they might drop the position prior to the Beijing ISG due to the opposition of many ARF states, but apparently had decided against this in the face of the PLA firsthand scrutiny of China's ARF diplomacy in Beijing.[52]

The evidence suggests, then, that over time the character of Chinese obstruction or resistance in its ARF diplomacy on the ground shifted from protecting given Chinese "interests" to, in part, protecting Chinese multilateral diplomacy from potential domestic opposition. Tentatively speak-

[50] The interviewee, in response to a question about the ACD Department's interest in Asia-Pacific multilateral security institutions, remarked that the department "had a big appetite" (January 12, 1999).

[51] Interview with senior US diplomat involved in ARF policy, February 2001; and interview with Canadian diplomats involved in ARF work, July 2001.

[52] Interview with Canadian diplomat involved in the ARF policy, May 5, 1997.

ing, one could plausibly see this as diplomacy more empathetic with the institution and less empathetic with other PRC constituencies that may have had different views of the value of multilateralism.

## Additional Empirical Implications

Even if the new multilateralism and mutual security discourse was not entirely cheap talk, has it been irrelevant talk in the sense that it has little constraining effect on behavior? This question is central, of course, for showing whether or how socialization matters. But it is not central in showing that socialization occurs. Policy outcomes, like international social structures, should also be seen as products of the interaction of multiple actors in bureaucratic social environments where persuasion, social influence, and mimicking, not to mention strategic behavior, may be at work. This is not the time nor the place to do this kind of analysis, though, of course, it needs to be done in order to show how all the components in figure 1.1 play out. Indeed, one has to have far more information about the highly secretive Chinese foreign policy process than is generally available now.[53] But a couple of points are worth mentioning.

---

[53] There is little doubt that the PLA has played a role in constraining the evolution of China's ARF policy. This is due to at least three factors. First, the military in general is more skeptical of multilateralism, as are militaries in many other nations. The CBM proposals that China has tabled which have raised the most opposition from other states in the ARF have tended to come from the PLA (e.g., observers at joint exercises; cessation of surveillance activities). This said, however, ARF diplomats have observed that the PLA officers most directly engaged in ARF diplomacy have been somewhat more supportive of multilateralism than other PLA officers (interview with Canadian diplomats involved in ARF diplomacy, July 2001; interview with PLA officer involved in ARF diplomacy, January 1999). Second, in the 1990s at least, policy coordination on the ARF was handled by an interagency process where the PLA appears to have had equal weight to the MOFA. ARF policy was coordinated by an Asia Pacific Security Coordinating Mechanism (*Yatai anquan xietiao* [sometimes *xieshang*] *jizhi*), nominally cochaired by a vice foreign minister and by the vice-chief of general staff in charge of PLA relations with the external world. Third, unlike the MOFA, the PLA did not have a dedicated functional office to handle or coordinate ARF policy. Rather it was handled by the Comprehensive Division of the Foreign Affairs Office of the General Staff Department. A senior officer in this office generally represented the PLA at ARF meetings. This office was understaffed, however, and could not do much original research on proposing or responding to ARF proposals. It therefore had to coordinate with other PLA institutions, such as the AMS Strategy Department, the National Defense University's Strategic Studies Institute, and the Second Department (Military Intelligence) of the General Staff Department. The process was cumbersome, and ad hoc, and other institutions grumbled that the FAO was unwilling to relinquish its control over ARF policy. The effect of these three factors was to give the PLA considerable influence on ARF policy, to slow down the PLA decision-making process on ARF issues, and to encourage a skeptical PLA eye on China's policy. This characterization of the previous role is based on interviews with a senior PLA officer involved in ARF policy, January 1999; with a senior analyst in a govern-

First, even though I am black-boxing part of the policy process (the part where the socialized protomultilateralists then interact with other constituencies and communities and their normative and causal arguments), suppose they had (and have) an influence substantially greater than zero, and suppose that the concept of mutual security had enough normative substance such that policy behavior ought to have reflected some of its elements. There ought, then, to be some empirical implications. First and foremost, one ought to see over time a greater degree of Chinese "comfort" even as the ARF became more institutionalized and developed a somewhat more intrusive agenda. That is, there ought to be things about the ARF that China later accepted that it either opposed (or counterfactually one could plausibly have expected China to oppose) just before it joined in 1994. In fact, the evidence does suggest that Chinese decision makers' changing comfort level allowed the following changes in the ARF institution and agenda.

*Institutional Structure*

The main innovation in the ARF structure occurred at the Second ARF in 1995. There the ARF agreed to set up two kinds of working groups to undertake intersessional discussions that could not be handled in the annual day-long foreign ministers' meeting. Canada and Australia had floated proposals at the First ARF in 1994 for Track I intersessional work, but these had been rejected at the time, primarily because of Chinese objections (Leifer 1996:32). In 1995 the proposal was put on the ARF agenda again. This time, despite some Chinese grumbling over the terminology and temporal mandate (China objected to the term "working groups" and to an indefinite time frame because both smacked of thicker institutionalization), the ARF created two intersessional meetings (ISM) (peacekeeping operations, and search and rescue) and one intersessional support group (ISG) on CBMs. In practice, these were similar types of activities. Their initial mandate was only to meet once in 1996, and the Third ARF would then decide whether or not to extend their lives, but they have been renewed regularly since then.[54] In 1998, the ARF ISG on CBMs recommended that the ARF convene two meetings of the ISG in 1999, further "regularizing" what is supposed to be

---

ment think tank involved in the ARF policy process, 1996; and with a senior analyst in a government think tank involved in Asia-Pacific security affairs, January 1999.

[54] Much of the previous paragraph came from an interview with a former senior US administration figure involved in Asia policy, Beijing, June 1996, from Smith 1997, and from an e-mail exchange with an Australian government official involved in ARF policy, January 1997.

an ad hoc process (ARF 1998a:8). The ISG and ISMs finally provided the ARF with a process for more detailed and sustained investigation of solutions to security problems in the region. This allowed states with particular expertise and/or interest to influence intersessional work (e.g., Canada and PKO). Most surprising to ARF participants, but consistent with the argument about China's increasing comfort levels, China offered at the 1996 ARF to cochair an ISG on CBMs with the Philippines in March 1997. China became a participant in the intersessional process in a way no one predicted in 1993.

*Agenda*

Here there have been a number of changes that were either rejected in 1993 and 1994, or were viewed as too controversial. All of these reflect some give by the Chinese. On nuclear testing, for example, despite its sensitivity to criticism on this score, the Chinese did not disrupt consensus when the 1995 and 1996 chair's statements indirectly criticized China (and France) for its nuclear testing programs (ARF 1995a:7, 1996c:3).

On preventive diplomacy (e.g., using the chair's good offices to investigate or mediate disputes, sending ARF special representatives on fact-finding missions, moral suasion, and third-party mediation) Chinese decision makers have traditionally been very uneasy with a more active ARF role because of the potential for the internationalization of core security issues such as Taiwan.[55] Nonetheless, the ARF formally took up the issue at its Track II working group on preventive diplomacy in November 1996 in Paris. Indeed, the explicit mandate of the Paris working group was to propose a list of relevant preventive diplomacy CBMs for the agenda of the ISG cohosted by China in March 1997. At the time, the main concrete recommendation to come out of the meeting was a proposal to expand the role of the ARF chair's "good offices" (ARF 1996c:2). In April 1998 at another ARF Track II working group on PD, the group agreed to recommend to the ARF SOM an "enhanced role for the ARF chair or other third parties in providing good offices in certain circumstances." This was a slight expansion in the list of whose good offices might be called upon for preventive diplomacy. Interestingly enough, China's own experience with border CBMs with the Indians and Russians was suggested as possibly relevant for PD in the rest of the region (ARF 1998b:5). These CBMs

---

[55] China was not alone. South Korea apparently was leery of giving the ARF a preventive diplomacy role if this meant that ASEAN might try to involve itself in Northeast Asian issues. Interview with prominent Canadian academic involved in Track II activities, January 1997.

were, on paper at least, "contractual" CSCE-like agreements placing specific limits on the size and movement of military forces along borders.

The issue moved from Track II to Track I at the Sixth Senior Official Meeting in May 1999, where it was agreed that at the CBM ISGs in 2000, the question of the ARF chair's good offices should be discussed in more detail. A draft paper on preventive diplomacy, prepared by Singapore, was circulated in November 1999 prior to the ISG on CBMs in Singapore in April 2000. The paper outlined the principles and scope of the concept.[56] The Singapore meeting authorized more explicit focus on an enhanced role for the chair and for "Experts/Eminent Persons" (EEP). Papers on these two topics, presented by Japan and Korea, respectively, were tabled later in 2000.[57] This finally initiated a detailed Track I debate in the ARF over PD.

The Chinese position evolved from opposition to PD to a more active, though still wary, diplomacy. The Chinese delegation officially contributed a working paper on PD in February 2000, prior to the Singapore ISG, in which it staked out key principles. These stressed that the ARF was a forum, not a mechanism "for dissolving specific conflicts" (China 2000a:2). Preventive diplomacy should use peaceful diplomatic means (by implication eschewing military operations such as PKO) to prevent armed conflict and only with the consent of all the parties directly involved. Any PD should also be based on mutual respect for sovereignty, territorial integrity and noninterference in internal affairs, and extant international law. On the basis of this paper, the Chinese suggested changes to the Singapore PD paper that would have, by and large, incorporated substantial portions directly from the Chinese working paper.

Some of these suggestions made their way into a revised Singapore paper in April 2000. Some of these changes were minor deletions of language. One, however, enshrined the principles of the UN Charter, the Five Principles of Peaceful Coexistence, and the ASEAN Treaty of Amity and Cooperation in one of the eight principles of PD, thus ensuring that the PD paper placed more emphasis on upholding sovereignty and noninterference in internal affairs. China (and other states) also beat back a Canadian effort to dilute this principle with language on respecting human rights and the rule of law. Not all of the Chinese suggestions were incorporated, in particular a proposal to delete language that, in its view, might

---

[56] See "ASEAN Regional Forum Concept and Principles of Preventive Diplomacy" (ASEAN Draft, November 6, 1999).

[57] "Enhanced Role of ARF Chair" (discussion of the ARF ISG on CBMs, 2000); "Co-chair's Draft Paper on the Terms of Reference for the ARF Experts/Eminent Persons (EEPs)" (October 2000).

allow PD in cases of bilateral disputes that had the potential of spreading to other states.[58]

The PD issue is, as of this writing, at a stage where states are agreeing to disagree about some of the principles and modalities of the PD. The revised Singapore paper was accepted at an ISG in Kuala Lumpur in April 2001, though as a "snapshot" of the state of discussions on PD and with acknowledgment that substantial differences remained on virtually all of its components. The fact remains, however, that the ARF appears still to be committed to developing a mechanism for more proactive interstate dispute prevention. Indeed, in 2005 it agreed to change the name of its ISG on CBMs to the ISG on CBMS and Preventive Diplomacy (ARF 2005). Chinese diplomacy on PD is no longer aimed at preventing this kind of evolution in the role of the ARF. Rather it has acquiesced to the notion of PD and instead has been essentially aimed at shielding the Taiwan issue and its own bilateral territorial disputes from ARF-based PD, and at strengthening language on sovereignty and noninterference in internal affairs.[59] The fact that the ARF took up these issues and is moving the discussion slowly forward despite China's concerns (and those of some other ASEAN states as well) suggests, again, a changing degree of Chinese comfort with the evolving agenda.[60]

On the South China Sea question, China's leaders' longtime preference has been for bilateral discussions with other claimants. They have worried that in multilateral settings China would be outvoted, its bargaining power diluted, leading to the dilution of China's sovereignty claims or, worse, the carving up of China's claims (Sun 1996:297). They tried assiduously in the past to prevent what they call the "internationalization" of the issue (Shang 1996:288–95). It was considered a conceptual break-

---

[58] This analysis is based on comparisons of the November 1999 version of the Singapore PD paper, the Chinese comments on this draft (China 2000b), the revised April 2000 version of the Singapore PD paper, and subsequent Chinese responses to this revised draft, submitted in January 2001 (China 2001).

[59] The language on sovereignty and noninterference was already quite strong in the original draft paper. That is to say, China was not an outlier in promoting this kind of language in ARF discourse. The outliers on PD have been the Canadians, who have pressed language that does not subject PD to consensus decision making, that includes respect for human rights, and that allows for military actions such as peacekeeping operations. Canadian diplomats have complained that many Western states have raised very few objections to the issues raised by the Chinese and others and have not been strong supporters of the Canadian position, fearing that the relatively fast-track intrusiveness of the Canadian proposals may undermine support for PD in the end. Interview with Canadian diplomats involved in ARF policy, July 2001.

[60] The MOFA-commissioned report on comprehensive security explicitly advocated strengthening the PD capabilities of the ARF, though there was no immediate concrete manifestation of this argument in China's diplomacy.

through, then, when the SCS was put on the Second ARF agenda in 1995. Even though internal reports indicated continuing fears of multilateral approaches to resolving the issue (Sun 1996:297), the Chinese delegation did not object to the chair's declaration of consensus. Nor was China willing (or able) to prevent the statement from pointedly encouraging all claimants to reaffirm their commitment to ASEAN's 1992 Declaration on the South China Sea—this after China's construction of a small naval post on the disputed Mischief Reef in February 1995. The Third ARF chair's statement in 1996 again touched on the SCS issue—this time welcoming China for its commitment in 1995 to resolve SCS disputes according to international law, but also pointedly commending the Indonesian workshop on the South China Sea for its work on conflict management issues (ARF 1996a:4). The workshop was set up in 1992 and is funded by Canada. The Chinese had been unhappy with this and had tried to pressure the Canadians to stop funding. By the Third ARF, apparently, China did not believe it was necessary to oppose consensus on this issue.

Finally, on CBMs, China was traditionally skeptical about the value of CBMs to the extent these are deemed asymmetrically intrusive. Weak states, like China, it claimed, should rightfully be less transparent than strong states like the United States. In addition, China has criticized the notion that one can transplant CSCE-type CBMs to the Asia Pacific. The First ARF was relatively silent on CBMs. However, by the Second ARF, under Brunei's leadership, the ARF had endorsed the ARF Concept Paper that laid out a timetable for implementing a wide variety of CBMs. These, all voluntary, would be taken from the Annex A list and included the following: statements on security perceptions and defense policies; enhanced military-to-military exchanges; observers at military exercises; promotion of the principles of the ASEAN TAC and the ASEAN Declaration on the South China Sea; exchanges of information on PKO activities, among others. At the ISG on CBMs in January 1996, states presented defense white papers and statements about security perceptions. But no comments on or criticisms of the content were permitted. There were complaints outside the ARF that the Chinese presentation—a white paper on arms control—was not especially detailed or credible. China followed up in 1998 with a more detailed and sophisticated white paper on defense, modeled more or less on the Japanese and British white papers.[61]

---

[61] Evidently the first white paper was called a white paper on "arms control and disarmament," even though it covered topics included in defense white papers, because top Chinese military leaders did not want it to appear that China was bowing to external pressure to produce a white paper on defense per se. Comments by senior National Defense University officer, March 1997. The drafters of the 1998 white paper explicitly examined a range of possible templates, and rejected some Southeast Asian examples for being too slim and nontransparent. Interview with Chinese military officer, 2000.

By the Third ARF, with the results in from the ISG on CBMs, the list of CBMs recommended in the chair's statement lengthened and deepened. Although defense white papers and statements on security policies were still voluntary, there were hints of an emerging template.[62] "Such papers could also cover defense contacts and exchange programmes undertaken by participants" (ARF 1996a:4). The statement also hinted that, unlike in the ISG, the content of these papers would also no longer be off limits to discussion. "Exchanges of views on the information provided in such statements and papers should be encouraged in future ARF dialogues" (ARF 1996a:5). On military observers at exercises and prior notification of military exercises, the statement noted that states were encouraged to exchange information about their ongoing observer and prior notification activities "with a view to discussing the possibilities of such measures in selected exercises" (ARF 1996a:6). The March 1997 ISG on CBMs cochaired by China and the Philippines pushed this further. The agenda for the meeting called for reaching consensus on the invitation of observers to joint military exercises and the prior notification of joint military exercises.[63] Interestingly, although ASEAN and China tend to decry the validity of a CSCE template for the Asia Pacific, the CBMs that are now either on the table in the ARF ISG or endorsed in the ARF Concept Paper Annex B (or embodied in the Five-Power Treaty) are not much different in kind from the first generation of CBMs under the CSCE (Desjardin 1996:7).

By the end of the decade, China had proposed or hosted a number of CBMs ranging from the Fourth APF Meeting of Heads of Defense Colleges, to a seminar on defense conversion cooperation, to exchanges on military law, to military exchanges on environmental protection.[64] The character of these proposals still reflected an impulse toward unilateralism—that is, they were all proposed by China without coordination with other states or without asking other states to cochair or co-organize. Moreover, some proposals were frustratingly vague. For instance,

---

[62] For which activists were pushing hard. See, for instance, CSCAP Memorandum No. 2, "Asia-Pacific Confidence and Security Building Measures" (n.d. ~ 1995), p. 4.

[63] The agenda focus on joint exercises had been set by China, and it ran into opposition from the United States and its allies, as they are the ones who run joint exercises. China does not. The ISG failed to reach consensus on the issue. Although this was a setback for the CBM process, a range of alternative proposals was floated for implementing other versions of this kind of CBM. The Chinese ran into heavy criticism for this proposal. Subsequently the proposal was dropped. Interview with Canadian diplomat involved in ARF policy, May 5, 1997; interview with PLA officer involved in ARF policy, January 13, 1999.

[64] See the Chairman's Summary, "The Fourth ARF Meeting of Heads of Defense Colleges" (Beijing, September 2000); "The ASEAN Regional Forum Seminar on Defense Conversion Cooperation" (Beijing, September 2000).

China proposed that a maritime information center be set up in Tianjin to provide the region with information about climate and ocean conditions, among other topics. Other delegates had a hard time trying to elicit more specific details about how such a center might be run and how the information might be disseminated (smaller states might be reluctant to rely on information controlled or provided by a great power in the region).[65] In addition, some of the CBM proposals are transparently self-serving, such as the previously discussed CBM on joint military exercises or a proposal for states to cease surveillance operations against each other. But the fact remains that this activity, while limited, was not viewed as particularly duplicitous by most ARF states, and was considered a welcome indication of a growing Chinese sophistication and nuanced commitment to multilateral measures.

Thus, change over time in the ARF was a result, in part, of the social effects of its initial form and function on one of the key actors in the institution. The mutual evolution between social environment and actor interests, understandings, and behavior is precisely what, according to constructivism, we should expect to see.

None of this means that China never got its way. Clearly, despite the changes, the institutionalization and agenda of the ARF did not move, and is not moving, as fast as some countries would like. But often the limits to Chinese comfort levels tended to show up in the language adopted, rather than in the concrete content of discussions. Although preventive diplomacy is on the agenda, the Chinese have been reluctant to support conflict resolution roles for the ARF. The 1995 ARF Concept Paper had divided the timeline for ARF development into three phases: CBM phase, development of preventive diplomacy phase, and a phase for the development of conflict resolution mechanisms. When the Second ARF chair's statement endorsed the Concept Paper, however, "conflict resolution" was changed to "the elaboration of approaches to conflict." The Chinese had objected to "conflict resolution mechanisms" because the term implied giving the ARF a mandate to intervene in conflicts that the Chinese might want to keep bilateral.[66] The slow pace of discussion on preventive diplomacy is, in part, a function of China's worries about its application to bilateral disputes, or conflicts it considers to be internal (e.g., Taiwan, ethnic separatism), though it has to be said that China is not alone in stressing the importance of the principle of sovereignty and independence in the application of PD mechanisms.[67]

---

[65] Interview with senior US diplomat involved in ARF policy, February 2001; and interview with Canadian diplomats involved in ARF diplomacy, July 2001.

[66] Interview with Singaporean embassy official, Beijing, April 1996; China 1996:3.

[67] Vietnam has been one of the more vocal opponents of efforts by countries such as Canada to dilute the sovereign-centric language or to include military options such as PKO

The second general point is that the multilateralism and mutual security discourse developed through involvement in the ARF and related activities may become even more constraining over time. Borrowing and modifying normative concepts are not cost-free. Alternative normative discourses can affect actor behavior in at least three ways. First, they can underscore a widening gap between discourse and practice.[68] Subjective pressure due to a perceived gap between one's new identity as embodied in the new discourse on the one hand, and identity-violating practices on the other, can lead to practices that are more consistent with the new identity (as consistency theory would suggest). Intersubjective pressure due to opprobrium generated when the new pro-group, pro-social discourse is obviously in tension with behavior can also lead to pro-group practices. In the China case, mutual security (at least its common security elements), the rejection of realpolitik *parabellum* (and incidentally a long-standing legitimate idiom in Chinese), and holding the Five-Power Treaty up as an exemplar of these principles, put China's violation of these principles in starker relief. This can have reputational costs in the contractual institutionalist sense, or legitimacy costs in domestic political processes, or opprobrium costs in terms of self-legitimation, identity consistency, and status.

Second, new normative discourses can positively sanction behavior that otherwise is unallowed or not seriously considered. For example, before China could enter international economic institutions such as the IMF and the World Bank in the early 1980s, it had to revise its long-standing support for Lenin's "inevitability of war" thesis. Early attempts in the late 1970s to do so ran into resistance.[69] Why? Because revisions would mobilize resistance from true believers, opponents of engagement with global capitalist institutions who could invoke Mao as legitimating their arguments. The "inevitability of war" discourse did not mean China was actually preparing for an inevitable global war between socialism and capitalism or within capitalism, nor that Chinese leaders necessarily believed it. But one could not be a Maoist and not believe it. Thus, one

---

in ARF documents on PD. See, for instance, "Vietnam's Views on Singapore's Discussion Paper" (January 2001); interview with Canadian diplomats involved in ARF diplomacy, July 3, 2001. India also stressed the importance of protecting sovereignty, and opposed military PD mechanisms. See its comments on the Singapore PD paper, "Preventive Diplomacy" (mimeograph, n.d.).

[68] See Keck and Sikkink's (1998) discussion of transnational activist and "accountability politics" and Kratochwil's discussion of Durkheim's analysis of morality (1984).

[69] This refers to the CCP's claim in the Maoist years that war between capitalism and communism was inevitable, and was the historical condition for the emergence of world socialism. Conversation with former senior Chinese diplomat and speechwriter for senior officials in the MOFA.

could not reject Maoism in foreign policy without rejecting the discourse. To no longer be a Maoist, to delegitimize Maoism as an obstacle to moving into institutions, required adjustment of the discourse. Revision of the discourse did not determine China's entry into the IMF and World Bank, but it permitted it, allowed action, and delegitimized Maoist opposition. In a similar fashion, mutual security and treating the Five-Power Treaty as an exemplar legitimated common security arguments internally. These also permitted proponents of these ideas to argue and defend their policies in ways that were illegitimate prior to China's entry into the ARF. That is, new discourses can legitimize or empower those who have genuinely internalized these norms to act politically, thus changing interagency balances of power and foreign policy outcomes.

Third, the logics and normative values embodied in discourses can constrain even those who use them instrumentally by narrowing the range of behavioral options that can be proposed or followed. It becomes harder for hyper-realpolitik actors to advocate unilateralist noncooperative security strategies if these fall outside of the range of behaviors acceptable in a cooperative security discourse.[70]

This does not guarantee the discourse will win out over realpolitik discourses and practices, and it does not mean there have not been other considerations behind the ARF policy process—image, rivalry with the United States, the mimicking of unfamiliar but standard diplomatic practices, among others. Nor does it mean there are no multilateral actions designed to enhance China's relative power in some way while diminishing the power of others. But it does suggest that by the late 1990s, there was now one more, legitimate, rival set of arguments that were normatively based on elements of common security and that committed China, perhaps unintended, to support more intrusive multilateral security measures that it would have opposed (indeed did oppose) prior to its entry into the ARF and associated regional security dialogues.

CONCLUSION

My tentative findings about the socialization of at least a portion of those Chinese officials exposed to the ARF have at least four objections. The first is that exogenous material side payments or threats may have been responsible for China's more constructive, comfortable approach to the ARF. This is fairly easy to handle: the ARF has no capacity to put any

---

[70] This is a point made in Evangelista's study of transnational effects on Soviet arms control policy. See Evangelista 1999.

such exogenous sanctions in place. Nor, as far as I am aware, did any other states unilaterally link any such sanctions to China's participation.

The second possibility is that changes in the nature of China's participation in the ARF over the 1990s reflected a deceptive effort to exploit cooperation from other states. One would expect this from a realpolitik actor with prisoners' dilemma preferences. However, if this were all only deception, we would expect that as the ARF handled increasingly intrusive and sensitive issues that could impinge on core interests or relative power issues, the PRC should balk at further change in the institution and agenda. In other words, the comfort level should be negatively related to the level of institutionalization and intrusiveness of the agenda. Yet, change in the ARF and change in comfort levels (at least of those participating in the ARF) were positively related over time. That the ARF moved from having a very innocuous agenda to discussions of more intrusive CBMs and preventive diplomacy mechanisms is evidence of this. Moreover, that eventually there were Chinese protomultilateralists who held up the Five-Power CBM Treaty as a potential model for East Asia suggests that prisoners' dilemma preferences were no longer uniform across the agents in the Chinese policy process.

The third objection is that China's participation and the overall rise in its multilateralist diplomacy in East Asia was and still is mainly aimed at countering US power and influence in the region. As Thomas Christensen argues, China's multilateralism is in part a (smart) reaction to a tougher US policy toward China in the second half of the 1990s (Christensen 2006:117–21). In other words, it is primarily a strategic response to US power, not a product of socialization processes. Or, at best, a tougher US policy was a precondition for the rise of multilateralist arguments inside the policy process. This argument has two problems. The first is, as I noted earlier, that some of the strongest advocates of the multilateralism discourse have tended not to be the same people or institutions as the more hard-line opponents of US military power in the region. Moreover, even if China's top decision makers (and their military advisers) have argued for using the ARF and other institutions as tools for constraining US military power, this does not mean that those exposed to the socialization processes in these institutions developed strategic or instrumental arguments in favor of multilateralism only.[71]

The second problem, however, is that both the shift toward a tough US policy and Chinese perceptions of this kind of shift occurred *after* the

---

[71] Socialization is about explaining how individuals change (or not), not about how the ship of state is reoriented (or not). Obviously, I believe the former is related to the latter, but to return to a point I made at the beginning of the book, the constructivist literature first has to show that the former does occur.

development of protomultilateralist arguments. Despite the 1996 Taiwan crisis, US China policy was even more clearly identified with an accommodationist constructive engagement from 1996 to 1999 than it had been prior. This period saw successful summit meetings between Jiang Zemin and William Clinton in 1997 and 1998. It also included more concrete US statements that US policy welcomed the emergence of a prosperous and stronger China. Finally, it also included much clearer US statements that the United States did not support Taiwan independence (Suettinger 2003:262, 284, 306, 348).

More important than US policy, however, were Chinese perceptions of this policy. In general, contemporaneous and subsequent analyses of US strategy toward China emphasize these positive developments. The turning point was in 1999, with the emergence of severe domestic criticisms of US engagement policy, the US bombing of the Chinese embassy in Belgrade, and increased tensions across the Taiwan Strait, among other events, which led to a serious downturn in the tenor of the relationship. Chinese analysts tend to define 1999 as the start of a more hostile US policy aimed, in their view, at a more vigorous containment of Chinese power in Asia (Li 2003:80–81, 90; Yan 2004:254–55, 257; Wu 2006:112–18; Fang 2003:299–300). There is no doubt that subsequently the proponents of a more proactive Chinese counter-containment strategy viewed multilateral diplomacy as one tool against US power, and particularly against closer US-Japan relations. But the origins of multilateralist ideas and their initial injection into the Chinese policy process preceded this perceived turning point in 1999.

The fourth objection is that China's changing comfort level is, in fact, a function of new information about the benign nature of the ARF. Beliefs about, and hence strategies toward, the ARF have changed, but preferences have not. This argument has a number of related components. The ARF has proven to be largely irrelevant to core security interests; most of the other participants have used the ARF to send assurance signals that it will not become an institution that constrains Chinese relative power; thus, the Chinese have discovered over time that it is relatively costless to participate in the ARF. There has been no real change in China's realpolitik, prisoners' dilemma preferences. At most, therefore, more cooperative behavior inside the ARF might serve short- to medium-term reputational purposes (in Kreps and Keohane's sense of reputation).

This last objection is the most serious and credible one. But I think it, too, has its problems. First, it is unlikely that a short-term concern for reputational benefits applicable to other specific opportunities for exchange was the driving force behind China's participation in the ARF. No other states, particularly those that could provide the most concrete costs or benefits to China—the United States and Japan, for example—were

linking ARF participation to other areas of cooperation such as trade. Indeed, the US government and Congress have been somewhat ambivalent about the value of the ARF. If the argument is that a more amorphous, long-term, instrumental notion of reputation mattered—that some material cost may be incurred, or some material benefit may be acquired somewhere in the indeterminate future from some other player(s)—then the reputational argument becomes virtually unfalsifiable.

Second, at the time China joined the ARF, it had not yet developed the overt strategy of improving economic and political relations with ASEAN that emerged in the late 1990s and into the early 2000s. This strategy was not formalized until the Sixteenth Party Congress in 2002, but the principles of such a strategy emerged around 1997 and the Asian financial crisis (Glosny 2007). There was no strong reason to believe that in 1994–1995, had China remained outside of the ARF, its economic relations with ASEAN would have been at risk.[72]

Third, from a theoretical perspective, the "new information" explanation at the heart of material reputational arguments is problematic because it underestimates the uncertain status of "new information." From the constructivist perspective, information is interpreted, and the same information can be interpreted differently in the context of similar institutional rules and structures. Empirically, we know that the same information will be interpreted differently depending on whether it comes from "people like us" (the information is more authoritative and persuasive) or comes from a devalued "other" (Kuklinski and Hurley 1996:127). Economic transactions—for instance, bargaining over price where people exchange information relating to their preferences and their "bottom line"—vary dramatically depending on whether or not the parties are friends (friends offer higher payments and lower prices than strangers) (Halpern 1997:835–68). Social context is an important variable in how well information reduces uncertainty in a transaction, and in which direction this uncertainty is reduced (e.g., clarifying the other as a friend or an adversary).

Thus, if all of China's ARF decision makers were realpolitik opportunists (that is, if they believed they were playing a prisoners' dilemma game in some form in East Asia) and if this basic worldview were fixed, then new information would be interpreted through these lenses. As I noted earlier, there is solid evidence from China's pronouncements and the interpretations of these by other states in the region that China initially looked upon multilateral institutions with a great deal of skepticism, and that its

---

[72] This, of course, does not mean that reputation is irrelevant for explaining the evolution of China's overall strategy toward ASEAN in recent years. See Glosny's (2007) detailed account of this strategy since 1997.

basic preferences were prisoners' dilemma ones. It is probably true that the initial signals provided by an underinstitutionalized and nonintrusive ARF in 1994 could have been interpreted as nonthreatening by realpoliticians.[73] But as the ARF agenda and institution evolved, the signals should have been interpreted with increasing alarm by realpoliticians, since the trend lines were toward issues and procedures that could place some limits on security policy options. Yet, for a small group of China's ARF policy makers, these signals were reinterpreted in less, not more, threatening ways. The fact that this group of policy makers and analysts eventually believed this information *was* reassuring while still expressing concern that others in the policy process (with more realpolitik views of multilateralism) might see this information as less reassuring suggests that the information provided by the ARF was often not unproblematically reassuring. Protomultilateralists did not enter the ARF with this more sanguine interpretation of "new information." Rather, this interpretation of the "new information" came from socialization inside the ARF.

I am the first to admit that the study of persuasion requires much additional fine-tuning of definitions, conditions, hypotheses, and analytical procedures. But if my general argument is plausible, then a focus on institutions as social environments raises important implications for institutional design. Typically, contractual institutions argue that the efficient institutional designs depend on the type of cooperation problem—e.g., a prisoners' dilemma–type problem requires information (monitoring) and sanctions; an assurance problem primarily requires reassurance information (Martin 1993). The flip side is that one can identify inefficient institutional designs for particular cooperation problems as well (e.g., an institution that is designed only to provide assurance information but has no monitoring or sanctioning capacity would be inefficient for resolving prisoners' dilemma–type problems). Additionally, Downs, Rocke, and Barsoom argue that so-called transformational institutions (inclusive institutions that bring genuine cooperators and potential defectors together in an effort to instill norms and obligations in the latter) are less likely to provide efficient solutions than a strategic construction approach. This latter approach to institutional design stresses exclusive memberships of true believers where decisions are made on the basis of supermajority rules. The gradual inclusion of potential defectors under these conditions ensures that the preferences of

---

[73] Though even this is problematic from an institutionalist perspective. As realpoliticians, the Chinese should have been especially suspicious of an institution that activist states such as Canada, Australia, and to some extent the United States supported. The information that their involvement should have supplied, for realpolitician skeptics, was precisely that the ARF was a potentially constraining institution.

the true believers predominate as the institution evolves. Their critique of the transformational approach rests explicitly on skepticism that the preferences of potential defectors can change through social interaction (Downs, Rocke, and Barsoom 1997, 1998).

It is not clear whether this skepticism rests on empirical evidence about the absence of state-level socialization, or simply, as I noted before, on the methodological difficulties of assuming and then trying to observe preference change.[74] In any event, *if* one relaxes the assumption about no preference change (particularly if, as I tried to argue here, this is based on theory and evidence of socialization in institutional environments), then one is forced to revisit the contractual institutionalists' notions of efficient institutional design. An institution that appears inefficient to contractual institutionalists (e.g., assurance institution for a prisoners' dilemma problem) may actually be efficient for the cooperation problem at hand. If, say, a player (or subactors in a policy process) with prisoners' dilemma preferences can be socialized to internalize assurance game preferences through interaction in a social environment with no material sanctioning or side payments, then "assurance" institutions may work in prisoners' dilemma–like cooperation problems. An efficient institution might then be reconceived as the *design* and *process* most likely to produce the most efficient environments for socializing actors in alternative definitions of interest.

---

[74] Indeed, the dependent variable in Downs, Rocke, and Barsoom (1997, 1998) is not the interests, preferences, or behavior of any individual state (say, the potentially worst or most important defector). Thus, it is possible that the (admittedly lower) observed level of cooperation (as they operationalize it) may still be a function of socialization. It is also possible, counterfactually, that the depth of cooperation might have been even shallower had there been no socialization effect in transformational institutions. Too much is black-boxed in their empirical tests, so they are unable to show one way or the other whether there are any socialization effects in either transformational or strategically and sequentially constructed institutions.