Decision Maker Preferences for International Legal Cooperation

Emilie M. Hafner-Burton, Brad L. LeVeck, David G. Victor, and James H. Fowler

Abstract Why do some decision makers prefer big multilateral agreements while others prefer cooperation in small clubs? Does enforcement encourage or deter institutional cooperation? We use experiments drawn from behavioral economics and cognitive psychology—along with a substantive survey focused on international trade—to illustrate how two behavioral traits (patience and strategic reasoning) of individuals who play key roles in negotiating and ratifying an international treaty shape their preferences for how treaties are designed and whether they are ratified. Patient subjects were more likely to prefer treaties with larger numbers of countries (and larger long-term benefits), as were subjects with the skill to anticipate how others will respond over multiple iterations of strategic games. The presence of an enforcement mechanism increased subjects' willingness to ratify treaties; however, strategic reasoning had double the effect of adding enforcement to a trade agreement: more strategic subjects were particularly likely to favor ratifying the agreement. We report these results for a sample of 509 university students and also show how similar patterns are revealed in a unique sample of ninety-two actual US policy elites. Under some conditions certain types of university student convenience samples can be useful for revealing elite-dominated policy preferences—different types of people in the same situation may prefer to approach decision-making tasks and reason through trade-offs in materially different ways.

What determines preferences for international legal cooperation? What determines how decision makers prefer to design or whether they aspire to ratify international

We thank the many participants who provided feedback on various drafts of this article at the International Studies Association, the American Political Science Association, the International Political Economy Society, the Burkle Center for International Relations at UCLA, the Department of Political Science at the University of Rochester, the Christopher H. Browne Center for International Politics at the University of Pennsylvania, the Mershon Center for International Security Studies at the University of Ohio, and the Department of Political Science at the University of Wisconsin. We are especially grateful to Avidit Acharya, Richard Herrmann, David Hsu, Robert Jervis, Patrick Kearney, Joshua Kertzer, Edward Mansfield, Helen Milner, Lisa Martin, Rose McDermott, Paul Poast, Kal Raustiala, Dustin Tingley, Alexander Thompson, Mike Tomz, Erik Voeten, and our anonymous reviewers. We are indebted to Alex Hughes and Linda Wong for invaluable research assistance, as well as Pamir Wang and Chris Clark for their assistance in conducting experiments at UCSD. We thank Manpreet Anand and Peter Cowhey for assistance in designing the survey instrument and David Robertson for assistance in building our sample. We thank the many participants who took the survey. The Laboratory on International Law and Regulation (ILAR) at UCSD funded this research. ILAR is supported by EPRI, BP plc, UCSD, and the Norwegian Research Foundation.

International Organization 68, Fall 2014, pp. 845–876
© The IO Foundation, 2014
doi:10.1017/S002081831400023X

https://doi.org/10.1017/S002081831400023X Published online by Cambridge University Press

treaties? Answers to these questions have focused on the types of problems that decision makers seek to manage through collective action,[1] domestic politics,[2] culture,[3] and the structure and cost of interstate bargaining.[4] Some also point to the costs of entry and accountability as selection mechanisms, suggesting that formal enforcement mechanisms diminish (or at times, augment) the will to join treaties.[5] Underlying much of this important scholarship are two assumptions: (1) that decision makers in the same situation will reason through the same trade-offs to arrive at similar conclusions about cooperation because context, structure, and institutions determine their preferences for cooperation; and (2) that different decision makers are similarly equipped to process large amounts of information, weigh costs and benefits far into the future, and act according to high levels of strategic reasoning.

Yet there are circumstances where individual political elites working within similar institutions, cultures, and contexts espouse radically different policy strategies. Writing in the middle 2000s, Carla Hills, the US Trade Representative under President George H.W. Bush, believed that the failure of the World Trade Organization's (WTO) Doha Round was attributable, in part, to lack of awareness by the US public of what was at stake for the country.[6] Susan Schwab, working in the same job for a president of the same party (the elder Bush's son, George W. Bush) and reflecting on the same challenge—whether or not (and how) to cooperate on Doha—at about the same time believed that Doha's troubles lay chiefly with the challenge of negotiating with more than 150 countries across so many issues.[7] From Hills's perspective, a strategy for crafting more effective trade agreements required a more active public relations effort at home; for Schwab, it required reducing the complexity of international bargaining. Why did two elites—each able to have substantial influence over what their government did in international trade relations—believe in such different policy strategies when so many of the other factors that characterized the situation were the same? One possible answer is that Hills and Schwab are very different types of decision makers, with different approaches to making the judgments needed to address the same policy challenges. This possibility arises not just with clearing the gridlock over Doha but in an array of pivotal decisions that lead to tangible outcomes in the design of international institutions and choices such as when to threaten and wage war.

The central hypothesis we explore in this article is that certain behavioral traits of the people tasked with making international cooperation decisions help to explain their preferences for negotiating and joining international agreements. We relax the assumptions widely used in standard international relations theories, allowing our

1. See Oye 1985; Stein 1982; and Sandler 2004.
2. Milner 1997.
3. See Meyer et al. 1997; and Boli and Thomas 1999.
4. See Keohane 2005; and Mearsheimer 1994.
5. See Downs, Rocke, and Barsoom 1996; and Von Stein 2005.
6. Hills 2005.
7. USTR 2008.

https://doi.org/10.1017/S002081831400023X Published online by Cambridge University Press

decision makers to vary in their behavioral traits. And we map that variation onto their stated preferences for cooperation in the context of an international trade agreement. Different types of people in the same situation may prefer to approach political decision-making tasks and reason through trade-offs in materially different ways.

We are hardly the first to suggest that behavioral traits could influence how people reason through political choices germane to international relations. It is well understood that people's attitudes, emotions, and even biology shape decisions that are typical of foreign policy.[8] Indeed, for years scholars of foreign policy have focused on individual elites—for example, Presidents John F. Kennedy's and Barack Obama's inexperience was thought to make each of them initially unable to form and implement foreign policy opinions independent of military advisers.[9] Scholars of presidential decision making have suggested that the methods for weighing evidence and making decisions adopted by individual leaders shaped how their administrations functioned, with a large impact on policy outcomes—including on matters of foreign policy.[10] In the realm of security studies, scholars have noted that traits such as overconfidence may make elite decision makers prone to error, leading to misperceptions, accidental wars, and other unfavorable foreign policy outcomes.[11] The study of behavioral psychology, however, has not made as much headway into the accepted canon of research on international institutions.

Our contribution is to measure, in a partially controlled setting, two key personality traits of decision makers—patience and strategic reasoning—and assess how subjects reason through foreign policy decisions related to trade cooperation. We do that with a large sample of university students, which affords the opportunity of statistical power, as well as a smaller sample of elite policy-makers, including former members of Congress, high-level political appointees and civil servants in the US government, and senior business strategists at leading American firms. By studying both populations using identical experimental instruments, we explore not just how behavioral traits relate to policy preferences for trade cooperation but also help reveal how nonelite samples may differ from real policy elites.[12] Our work thus offers a benchmark for explaining how individual behavioral traits might influence decision makers' preferences for treaty design and participation. It also offers a

8. See Tomz 2004 and 2008; Fowler and Schreiber 2008; Hafner-Burton, Hughes, and Victor 2013; Byman and Pollack 2001; McDermott 2008; and Hatemi and McDermott 2012.

9. Schlesinger 1965, chap. 10; on Obama, see Woodward 2010. Other work on how traits of leaders affect policy includes notably McDermott 2008.

10. For example, Neustadt 1960; and Greenstein 2004.

11. See Jervis 1968 and 1976; and Johnson 2004.

12. There is a growing and promising literature that uses survey experiments to probe how individuals make decisions related to international relations; see Tomz 2004 and 2008; Putnam and Shapiro 2009; Hainmueller and Hiscox 2010; Gartner 2011; Grieco et al. 2011; McDermott 2011; Mintz, Yang, and McDermott 2011; Tingley and Walter 2011a and 2011b; Tingley and Wang 2010; Tingley 2011; and Trager and Vavreck 2011. So far, none of this experimental research has looked at how behavioral traits affect preferences.

https://doi.org/10.1017/S002081831400023X Published online by Cambridge University Press

framework for exploring when and how nonelite samples might reveal insights about elite decision making relevant to international relations.[13]

First, we develop the theoretical intuition that measured levels of patience and strategic skills might influence how a decision maker reasons through particular treaty negotiation and ratification decisions. After outlining the study design, we report results from the university student sample. The higher statistical power of that sample allows us to show with reasonable confidence that our subjects' preferences regarding treaties depended—in part and possibly a lot—on their traits, quite apart from the context of the problems they were charged with solving. Behavioral traits may also have an effect on preferences of the same magnitude as some of the factors that have dominated international relations theory such as the presence of formal enforcement mechanisms—an institutional design feature that has been prominent in scholarly debates over the membership and effect of international agreements. We then turn to the elite sample, which is smaller but allows us to see if the patterns observed with nonelites are also evident in a population of people who actually reason through similar policy decisions for a living. Finally, we explore some of this work's implications for the study of international relations, including the behavior of agents in international negotiations and the question of where individual decision makers can inform real-world policy decisions in international relations.

How Behavioral Traits May Shape Preferences for Cooperation

Much of formal international cooperation starts with the design of international agreements that states join through ratification. Designing and joining an international treaty depends on a large number of decisions that come with trade-offs. We focus on the trade-offs involved with two decisions that correspond with topics that have attracted special attention by scholars of international organization: bargaining complexity and enforcement.[14] In addition to scholarly interest, policy-makers also behave as if these actions are important—something revealed by the huge literatures on how these factors affect the ratification process.[15]

The first decision we investigate involves how many parties decision makers prefer to invite to the negotiation. Scholars have explored how the number of parties in a negotiation affects the complexity of bargaining and the prospects for cooperation.[16]

13. For elite and nonelite comparisons, see Mintz, Redd, and Vedlitz 2006; as well as Helland, Loyning, and Monkerud 2013. The Pew Research center has also conducted a number of surveys exploring elite and nonelite opinion on foreign policy issues. For an example, see Gallup Organization and Times Mirror Company 1989; for studies that have compared populations on the trait of "loss aversion," see List 2003; and List and Mason 2009; regarding "trust," see Hedinger and Götte 2006; on "frames of reasoning," see Cooper et al. 1999; on styles of reasoning (systemic versus automatic), see Ackerman and Schneider 1985.

14. See Smith 2000; Raustiala 2005; Mansfield and Pevehouse 2008; and Hug and König 2002.

15. For example, Card et al. 2011.

16. See Kahler 1992; Alter and Meunier 2009; and Keohane and Victor 2011.

https://doi.org/10.1017/S002081831400023X Published online by Cambridge University Press

To measure how decision makers in our study managed a common trade-off involving the numbers of participants, we asked subjects to choose how many countries should be involved in the negotiation of a trade agreement. The survey included explicit instructions noting that while adding more countries to the negotiations would further their country's aim of having the agreement cover the largest fraction of world trade, the extra voices would also make bargaining more complicated and introduce additional risks that the content of the agreement would be diluted.[17] Responses were recorded in a category from 1 to 4, with each category increasing the number of countries invited to negotiate the treaty.[18]

The second choice we investigate involves the decision to ratify a trade agreement. A long-standing, central concern for scholars of institutions is whether and how international cooperation requires formal enforcement mechanisms such as dispute resolution procedures, and whether those mechanisms encourage or deter participation.[19] To measure how the presence of enforcement affected subjects' willingness to join a trade agreement, we randomly assigned subjects to one of two experimental conditions. In each condition the description of the trade agreement was the same except for one sentence that indicated whether the treaty included a formal enforcement mechanism.[20] We designed the survey to trigger subjects to weigh alternative policy options with the goal of maximizing the net national benefits from the trade agreement. We explicitly framed the decision to join the treaty as a situation where joining is a strategic complement. That is, states benefit more from joining and abiding by the agreement if other countries do the same. However, we did not specify just how much other states would actually benefit or whether some states might still have a dominant incentive to violate the agreement. In this type of strategic situation, cooperation can still be an equilibrium strategy even without enforcement, but enforcement makes it more certain that no state has a dominant strategy to free-ride—this additional certainty makes joining more attractive.[21]

Our central hypothesis is that certain behavioral traits of the decision makers tasked with making international cooperation choices help to explain their preferences for both of these decisions. The burgeoning research in experimental psychology and behavioral economics shows that people have many distinct behavioral traits (or

17. Although there are other ways to structure the potential trade-offs involved in this decision, this particular trade-off is commonly assumed in the literature.

18. Full information is provided in Section B of the online appendix.

19. See Fearon 1998; von Stein 2005; Gilligan 2006; and Donno 2010.

20. Subjects in the enforcement condition were told: "An independent enforcement mechanism promptly and credibly punishes any country that does not comply by taking away some of the benefits of the treaty from the country that breaks the rules." Subjects in the nonenforcement condition were told: "The treaty does not provide any formal mechanism to punish countries that fail to comply."

21. This type of situation is sometimes modeled formally as a global game, where cooperative actions are strategic complements, but where actors lack common knowledge of the payoff to cooperation and may even have different payoffs. Typically actors in these games play a threshold strategy, where they cooperate if their signal about the payoff is high enough and if their uncertainty about others' incentive to defect is low enough. For an overview of such games, see Myatt, Shin, and Wallace 2002.

https://doi.org/10.1017/S002081831400023X Published online by Cambridge University Press

dispositions), some of which can drive their social and strategic performance.[22] By "behavioral trait," we are simply referring to a behavioral tendency that, at a given point in time, is stable across similar types of decisions (in contrast to a state that often refers to emotional or other fleeting responses to a situation). For example, people who are patient are more willing to wait for greater rewards across a number of scenarios.

Scholars in the fields of cognitive psychology and behavioral economics have developed techniques for measuring a large number of traits and have explored a myriad of ways these traits might associate with individual and collective behavior. Eighteen traits in particular have commanded substantial empirical research.[23] Several of those traits have already spread into the American politics subfield of political science where scholars have probed how "agreeableness" or "conscientiousness" affect the mass public's political attitudes and civic engagement.[24] In principle, any trait that affects perception and reason might, in turn, affect how decision makers reason through trade-offs in international relations. Although such possibilities are interesting and could frame a research agenda for the long term, there is almost no research (to our knowledge) linking these measured behavioral traits to the reasoning and decisions typical of foreign policy. Thus, as a starting point, we have focused on traits that are highly likely to have a measurable impact on how decision makers choose among the trade-offs that are typical of one of the most central areas of foreign policy—international trade—where important decisions revolve around the design and adoption of international legal agreements. We focus on two traits—patience and strategic reasoning. We reveal how these traits affect outcomes but future research might explore other traits as well.

First is patience. International relations theorists have long known that one of the key functions international institutions perform is to lengthen the shadow of the future.[25] Institutions facilitate reciprocity and provide information that helps players become confident that the long-term benefits from repeated interaction will arise. Indeed, while international relations theorists disagree about a lot, this cooperation-enhancing role of the shadow of the future is one theory that commands widespread acceptance. At the level of individual elites who make policy decisions, patience is a trait that has the potential to affect the shadow's length. Patient people have lower discount rates—they are more willing to wait for larger benefits that accrue in the distant future rather than seize smaller but more immediate gains.

International cooperation frequently involves sacrificing payoffs that are immediate for benefits that are delayed. Faced with that particular trade-off, patience could

22. See Neale and Bazerman 1985; Knetsch 1989; Chen and Chaiken 1999; Costa-Gomes and Zauner 2001; Camerer, Ho, and Chong 2003; Fehr and List 2004; Fowler and Schreiber 2008; Tingley and Wang 2010; and Tingley 2011.
23. Hafner-Burton, Hughes, and Victor 2013.
24. For a review, see Gerber et al. 2010. See also Mondak and Halperin 2008; Vecchione and Caprara 2009; and Mondak et al. 2010.
25. Axelrod and Keohane 1985.

https://doi.org/10.1017/S002081831400023X Published online by Cambridge University Press

affect a person's preferences related to international cooperation in at least two ways. First, patience could affect how a decision maker prefers to pursue treaty negotiations when tasked with negotiating a treaty. Different bargaining strategies can substantially influence the length of negotiations and their prospects for success. For instance, including more countries in negotiations could potentially bring greater benefits in the long run because agreements would engage a larger share of the world economy; patient people are less concerned by the risk that negotiations might drag on and delay the benefits of cooperation.[26] Thus, our first hypothesis is that patient people in our study will be more willing to prefer complex negotiations, involving more countries and higher payoffs, even though they are made aware that doing so may delay the completion or scope of negotiations. Less patient subjects will prefer to invite fewer parties to the bargain.

Second, when tasked with the decision whether to join an international treaty, a decision maker's patience could affect whether s/he views the commitment favorably, especially in situations of time inconsistency when participation within the institution entails proximate costs but yields the possibility of large yet more remote benefits.[27] We anticipate that the type of person who is generally willing to wait for higher payoffs—a patient person—will prefer to engage in international cooperation, and join the treaty, more so than the impatient type who seeks more immediate gratification.

Our arguments about patience, applied in this study for the first time (to our knowledge) to international treaty cooperation, have a strong counterpart in formal models of international bargaining. Powell noted that many bargaining models are sensitive to assumptions about how much states value future payments.[28] More patient states are more willing to bargain for longer periods of time to secure peace rather than immediately secure a less valuable outside option. Leventoğlu and Tarar formalized this argument, showing that whether or not a negotiated settlement is attainable in many models of bargaining under incomplete information depends on the bargainers' patience.[29] These insights could also inform decision makers' preferences for cooperation where the rewards are more favorable trade agreements instead of the peaceful division of a prize.

Strategic reasoning is another behavioral trait that could influence how a decision maker reasons through these particular trade-offs. Economic theory distinguishes between choices that are game theoretic (that is, strategic) versus those that are simply decision theoretic. Strategic decisions are characterized by multiple decision makers, each faced with choices whose consequences depend also on others' choices. By contrast, decision-theoretic problems may depend on variables that are uncertain but they do not depend on the choices of other decision makers. Obviously, this

26. Martin 1995.
27. For example, Fowler and Kam 2006.
28. Powell 1999.
29. Leventoğlu and Tarar 2008.

https://doi.org/10.1017/S002081831400023X Published online by Cambridge University Press

distinction refers to ideal types, since many decisions combine game-theoretic and decision-theoretic elements. In this article we focus on game-theoretic choices because they usually require actors to form a clear and accurate picture of other people's incentives and choices and because many aspects of international cooperation have game-theoretic attributes.

For decades, scholars have known that most areas of international relations are steeped in strategy.[30] How a state behaves depends on what it expects other countries to do and on the vulnerability that each state has to others defecting. The application of these insights, which has been the province of game theory,[31] depends in part on the strategic reasoning of decision makers—that is, the extent to which they anticipate how their counterparts will respond in a bargaining situation and adjust their own response accordingly. Research in behavioral game theory suggests that there is wide variation in how people respond to strategic problems.[32] This is especially true when people initially play a game.[33] In these situations, individuals have not had time to form clear expectations for how others will play the game, and often make decisions about their own best move using simple nonequilibrium models of other players.[34] For example, some individuals may simply model their opponent as an actor who chooses randomly, and others may act as if their opponent always best-responds to random play, even when doing so is not an equilibrium strategy.

Given the central role for strategic interaction at the international level, strategic reasoning is a logical place to observe whether individual behavioral traits could affect policy preferences for international trade cooperation. This heterogeneity in strategic reasoning could have substantial implications for how people reason through the value of different forms and levels of international cooperation. Many elements thought to be important to the design and operation of international law—such as reciprocity—rely on the assumption that decision makers engage in high levels of iterated, strategic thinking. The preference to ratify and thus be bound by an agreement may depend partly on whether other states will also join and comply; policy-makers in those others states, as well, face a similar interdependent choice. Decision makers who treat the decisions of others as independent or random may be less likely to prefer joining an agreement because they do not fully consider the benefits from cooperation due to reciprocity. By contrast, those who realize that their decision to join an agreement will amplify the incentives for other countries to do the same may be more favorable toward joining. In fact, Jervis has pointed out that a key factor underlying uncooperative behavior during the Cold War may have

30. Stein 1982.
31. For a recent review of political science research on international legal institutions, including empirical work drawn from game theory, see Hafner-Burton, Victor, and Lupu 2012. For a more general review of game theory's role in models of international institutions, see Gilligan and Johns 2012.
32. See Nagel 1995; Stahl and Wilson 1995; Bosch-Domènech et al. 2002; Camerer 2003; Camerer, Ho, and Chong 2004; and Costa-Gomes and Crawford 2006.
33. Crawford, Costa-Gomes, and Iriberri 2013.
34. Stahl and Wilson 1995.

https://doi.org/10.1017/S002081831400023X Published online by Cambridge University Press

stemmed from individual decision makers failing to anticipate how others would respond to their own actions.[35] A similar idea has been posited by recent studies in neuro-economics that show that more strategic reasoners are more cooperative in iterated assurance games because they better anticipate other players' incentives to cooperate in response to their own cooperation.[36] Thus, our third hypothesis is that we expect people whose behavioral traits include higher levels of strategic reasoning will be more favorable about joining a treaty.

It can also be difficult to anticipate and best respond to the actions of a large and diverse group. Strategic actors especially may be more willing to engage in more complex negotiations than less strategic actors. Compared with nonstrategic actors who act without regard to what others will do, strategic decision makers act as if they are also better able to foresee and best respond to the actions of others. They may therefore be more confident in their ability to navigate negotiations involving a wider array of actors. Thus, we also expect subjects whose behavioral traits include deeper levels of strategic reasoning will prefer more complex treaty negotiations—willing to negotiate with more countries whose interests must be represented and whose actions must be predicted. Table 1 summarizes our argument.

TABLE 1. *Hypotheses*

	Preferences for outcomes	
Traits	*Negotiation strategy*	*Cooperation*
Patience	More patient decision makers are more willing to endure complex negotiations (with many more parties) to reap higher gains.	More patient decision makers will be more willing to join treaties with delayed benefits.
Strategic reasoning (Level-K)	More strategic decision makers are more confident about their ability to successfully navigate complex negotiations and thus are more willing to engage in negotiations with many parties.	More strategic decision makers are more willing to join cooperative agreements because they understand that their joining affects whether other states will reciprocate.

Study Design

Because we aim to explore the relationship between two behavioral traits and structured choices about international treaty cooperation in different situations we asked both our university and elite subject pools to participate in a survey experiment in two parts. One part asked them to self-report how they would respond to different scenarios and choices about an international trade agreement given the specific

35. Jervis 1998, 258–60.
36. See Yoshida, Seymour, Friston, and Dolan 2010 and Yoshida, Dziobek, et al. 2010; and Yoshida, Dolan, and Friston 2008.

https://doi.org/10.1017/S002081831400023X Published online by Cambridge University Press

trade-offs we outlined earlier.[37] The other part asked subjects to play a battery of behavioral economic games from which we elicited information about behavioral traits. We also asked a standard set of demographic questions that collect information about age, sex, and political party identifiers.

Both the survey questions and behavioral games in our university study were administered as part of a larger omnibus study where subjects participated in a number of short surveys and experimental tasks contributed by different researchers. The order in which subjects participated in each task was randomized to avoid any potential order effects. We report the questions and games relevant to the current article only.[38]

We conducted our university study during fall 2010 and winter 2011 at the Rady School of Business's behavioral computer lab, located on the University of California San Diego (UCSD) campus. A total of 509 participants were recruited from university classes in the Political Science Department, as well as two first-year masters courses at UCSD's School of International Relations and Pacific Studies. Students were compensated for participating in the study by receiving extra credit in the class from which they were recruited. All participants also had the chance to win monetary rewards from a lottery whose value depended on how they and other respondents played the experimental games in the study. Instructions for both the games and the survey were presented to subjects on computers in the lab using Qualtrics survey software. The entire enterprise was approved and overseen by UCSD's Human Research Protections Program.

Measuring Behavioral Traits

To measure the two behavioral traits studied here we rely on behavioral games rather than traditional survey-based measures for two reasons. First, unlike traditional surveys that ask subjects to categorize their own traits—say, by rating their own level of patience—behavioral games are based on a mathematical model of economic behavior, giving researchers a common baseline against which to compare subjects' behavior.[39]

Second, unlike traditional surveys, the games used in this survey force subjects to make decisions linked to tangible outcomes (real monetary stakes) and thus are probably a more accurate elicitation of underlying traits. This is especially important when studying factors like patience that are considered socially desirable and thus may be

37. We cannot guarantee that the answers correspond to what subjects would actually do in these situations, although they have no clear incentive to lie. It is, of course, difficult to know how sensitive their preferences are to the particulars of our scenarios. However, in constructing our survey we consulted with actual trade negotiators to make sure the choices reflected real trade-offs. Thus, we can be relatively confident that responses are informative about a real and important subset of treaty negotiation scenarios. Furthermore, the people we consulted agreed that there was not an obvious answer to these questions, which made it less likely that responses would be contaminated by a desirability bias, where subjects try to give the response that they think "looks good."

38. A full list of the tasks in which the subjects participated is available from the authors on request.

39. Camerer 2003.

https://doi.org/10.1017/S002081831400023X Published online by Cambridge University Press

prone to biased self-reporting.[40] The average monetary stakes that subjects faced in our experiments were quite small because a lottery (for $100) paid only a small number of subjects. However, other studies have shown that similarly small incentives are adequate for eliciting accurate responses,[41] even when decisions were elicited with a single prize.[42] Furthermore, a number of studies have reliably found a link between behavior in lottery-incentivized games and real-world political behavior.[43]

Patience. To measure how much subjects value the future—that is, their level of *patience*—we adapted a "choice game" introduced by Coller and Williams.[44] We refer to this game as a time-discounting task to more intuitively evoke the game's purpose. Past studies have linked behavior in this task to real-world behavior, such as savings rates[45] and voting behavior.[46]

In our study (and in others using this task), our subjects were asked to make twenty different choices between a $100 prize that would be paid to them within thirty days after taking the study and a variable, larger prize that would be paid within sixty days. For each subject, a measure of patience is the number of sixty-day choices. (Additional discussion of how these choices relate to discount rates is in the online appendix.)

Figure 1 shows the distribution of time-discounting choices made by subjects in our study, which is similar to choices found in many other studies.[47] The modes at the extremes indicate that many subjects either always chose one option or always chose the other. Heuristics also generated some modes—for instance, the large spike at five corresponds to subjects switching when the sixty-day choice moves above $110—a result similar to that in other studies.[48]

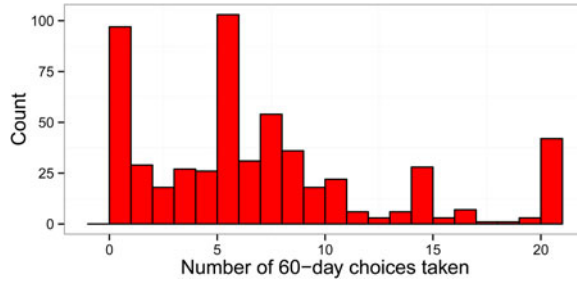


FIGURE 1. *Distribution of patience in the college student sample*

40. See Mischel 1974; Kagan 1988; Schwarz 1999; Berinsky 2004; and Webb et al. 2000.
41. See Camerer and Hogarth 1999; and Palfrey and Wang 2009.
42. Coller and Williams 1999.
43. See Fowler 2006; Fowler and Kam 2006 and 2007; and Dawes, Loewen, and Fowler 2011.
44. Coller and Williams 1999.
45. Harrison, Lau, and Williams 2002.
46. Fowler and Kam 2006.
47. See ibid.; and Coller and Williams 1999.
48. Notably see Harrison, Lau, and Williams 2002; and Fowler and Kam 2006.

https://doi.org/10.1017/S002081831400023X Published online by Cambridge University Press

Strategic reasoning. The game most frequently used to study a subject's depth of reasoning in games is the beauty contest, originally implemented by Nagel.[49] In this game, *N* players are asked to pick a whole number between 0 and 100 (inclusive). The winner of the game is the player who picks a number closest to the population average multiplied by a number *M*. If *M* is less than 1, the unique equilibrium strategy is for all players to guess 0; when *M* exceeds 1, the unique equilibrium is for all players to guess 100. For example, imagine a version of the game where the multiplier is 2/3. A player starts with a conjecture that other players choose numbers such that the average is 50. The player should then select 33 as his own choice since this is the closest number to 2/3's the group average. Other rational players know this; they, too, choose 33. But if everyone chooses 33, then the original player's best pick is 22 (2/3 of 33). At infinite iteration the best choice is 0. To characterize the number of rounds of strategic reasoning used by subjects we rely on Stahl and Wilson's "Level-K" model. In this model Level-0 players are nonstrategic; they play a random strategy. Level-1 players best respond to average Level-0 play by picking 50 × M—a choice that reflects one round of iteration. Level-2 players best respond to Level-1 players by picking 50 × M^2—two rounds of iteration—and so forth. Although it is possible for players to iterate to an infinite number of levels, most strategies are found to correspond to 1 or 2 steps of iterated reasoning at most.[50]

Level-K measurements are not just an assessment of the player's skill at iterative strategic thinking but also their expectation of what others will do when facing the same choices.[51] If Henry Kissinger thought all other players were mere randomizers then his best choices would be Level-1, but if he thought his opponents were Kissingerian in strategic skill his Level-K would be much higher. Thus, we measure how strategically individuals tend to act, not necessarily their strategic capability.

Most importantly, however, our results do not strongly depend on particular assumptions about why certain individuals tend to act more or less strategically. Our key assumption is only that this tendency is partially preserved across different decisions and this assumption has been supported by recent studies that show how subjects' level of strategic reasoning frequently persists across different types of strategic interactions.[52] Thus, while our measure may reflect both subjects' cognitive ability, as well as their beliefs about other players, it is also quite plausible that players who act more strategically in the beauty-contest games we employ also act more strategically in other decision-making domains, such as decisions about international cooperation.[53] This is especially likely given the finding that subjects'

49. Nagel 1995.
50. See Nagel 1995; Bosch-Domènech et al. 2002; Camerer 2003; and Costa-Gomes and Crawford 2006.
51. See Agranov et al. 2012; and Le Coq and Sturluson 2012.
52. See De Sousa, Hallard, and Terracol 2012; Bhui and Camerer 2011; and Agranov, Caplin, Tergiman 2013.
53. One way our results could be affected for our university sample is if university subjects acted less strategically in our games because they were playing other students, but more strategically in the

https://doi.org/10.1017/S002081831400023X Published online by Cambridge University Press

measured Level-K behavior is particularly stable when it is elicited through multiple games,[54] which is the method employed here.

In order to identify the level of reasoning typically employed by a subject in our study, we followed Coricelli and Nagel by having each subject play multiple beauty-contest games, each with a different multiplier.[55] Unlike single games, this approach creates a fingerprint that better identifies a player's typical depth of reasoning in the game. Figure 2 shows the patterns of choices that Levels 0, 1, and 2 would make across each of the six games. As one can see, Level-0 players should consistently play a random strategy across all games, on average choosing 50. Level-1 players best respond to this by multiplying 50 by the game's multiplier, creating a linear strategy profile across games. Level-2 and higher players will play numbers that are nonlinear in the multiplier.

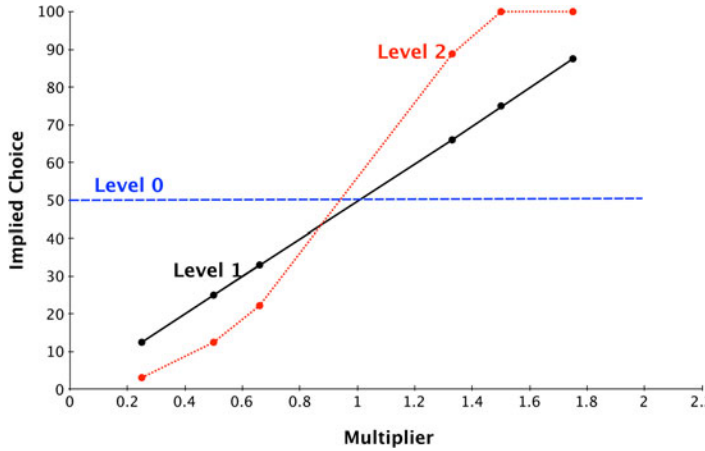


FIGURE 2. *Expected Level-K responses across six beauty-contest games*

University subjects were told that we would pick one game at the end of the academic quarter, and pay the winner of that game $100. Elite subjects were told that a winner would be chosen and paid at the conclusion of the study. In both cases it was made clear to subjects that they were playing against their peers. Following Coricelli and Nagel, for every choice in every game we calculated which level had the lowest

negotiation decisions because they thought they were (hypothetically) facing more strategic actors. If this were true it would weaken the link between strategic behavior in the beauty-contest game and responses to negotiating decisions. Thus, we risk underestimating the link between strategic ability and international decision making. However, we have no evidence that this occurred. As we show later, the effect of Level-K behavior is relatively similar for elites (who knowingly played other elites in the beauty-contest game and in our vignettes) and undergraduates.

54. Bhui and Camerer 2011.

55. The multipliers we used were 1/4, 2/3, 1/2, 3/2, 4/3, and 7/4. The unique equilibrium strategy for all multipliers less than or equal to 1 is 0. For all multipliers greater than 1, the equilibrium strategy is 100; Coricelli and Nagel 2009.

https://doi.org/10.1017/S002081831400023X Published online by Cambridge University Press

squared error. We labeled players as Level-2 if they had four or more of their six choices that were closest to the choice played by an archetypical Level-2 player such as in Figure 2.[56] As with Coricelli and Nagel, we do not calculate levels higher than 2. Thus Level-2 players in our data actually represent "Level-2 or higher." We labeled players as Level-1 if four or more of their choices were closest to a Level-1 or greater. We labeled remaining players as a Level-0.

Figure 3 shows the frequency of subjects estimated to be in each category. Most subjects in our study are either categorized as Level-0 players, who act randomly with respect to their payoffs or Level-1 players who best respond to random play. A small fraction is categorized as Level-2 or higher. The relative size of each group is consistent with what Bosch-Domènech and colleagues found for experiments in classrooms and laboratory settings.[57] In sum, only a few of these subjects act in a highly strategic manner.

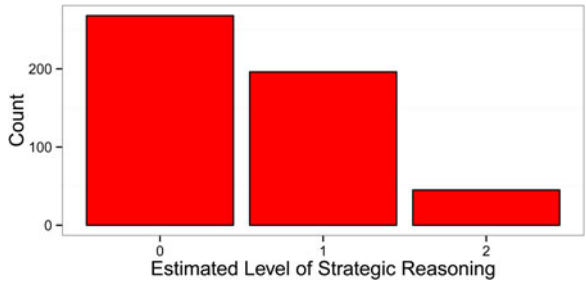


FIGURE 3. *Distribution of Level-K reasoning in the college student sample*

Although a number of studies have used the beauty-contest game to study strategic reasoning, we believe we are the first to try and connect results in this game to behavior in a separate decision-making domain, such as the design and joining of international agreements.

Results

Negotiation Complexity

We now look at how our sample of university subjects' patience and strategic reasoning are related to their preference for negotiating with more countries.[58] Table 2 reports three

56. Coricelli and Nagel had subjects play twelve games and categorized subjects as Level-2 if more than half of the subjects' decision (seven out of the twelve) corresponded to a Level-2 or greater player.
57. Bosch-Domènech et al. 2002.
58. Our hypothesized effects of patience and strategic reasoning are contingent on decision makers' strategic objectives given the particular set of trade-offs we specified in the survey. Different objectives or trade-offs could produce different hypotheses about the ways in which these traits affect preferences.

https://doi.org/10.1017/S002081831400023X Published online by Cambridge University Press

linear models, each regressing subjects' responses (1-4) onto our measures of behavioral traits.[59] The first set of results regresses subjects' choices on their measured level of patience.[60] The second set regresses subjects' choices on their measured level of strategic thinking, with Level-0 thinkers as the omitted category. The third set reports a regression that includes both measures, plus a number of standard demographic controls including the subject's years of postsecondary education (ranging from 1 to 7), gender (coded 1 for female), family income in the year before they entered college (coded on a scale from 1 to 9, with each number representing an interval of income),[61] and a dummy variable indicating subjects who were in graduate school. The third regression includes fewer subjects because some subjects did not report their family's income.

TABLE 2. *Behavioral traits and tolerance for complexity in negotiations*

	Dependent variable: Number of countries in negotiation (four categories)		
	Patience	*Level-K*	*Both measures + controls*
PATIENCE	0.016***		0.017***
	(0.004)		(0.004)
LEVEL-1 REASONER			
		0.072	0.043
		(0.048)	(0.051)
LEVEL-2 REASONER		0.281***	0.297***
		(0.099)	(0.105)
FEMALE			0.127***
			(0.048)
INCOME			0.006
			(0.012)
YEAR IN SCHOOL			−0.019
			(0.019)
GRADUATE STUDENT			−0.088
			(0.091)
INTERCEPT	1.974***	2.030***	1.904***
	(0.034)	(0.035)	(0.082)
R^2	0.01	0.01	0.03
N	509	509	487[1]

Notes: Standard errors in parentheses are White's robust standard errors. *** $p < .01$; ** $p < .05$; * $p < .1$.
1. This regression has fewer observations because some subjects did not report their family's income.

We find that both patience and strategic thinking are positively and significantly related to the number of countries a subject decides to invite to the negotiations. Subjects who made more sixty-day choices in our time-discounting task invited

59. For all regressions reported, we also ran ordered probit models, which relax the assumption that the four categories are equally spaced. These models yield substantively similar results and are available from the authors on request.
60. This is the number of sixty-day choices subjects made in our time-discounting task. See the section on measuring behavioral traits for further details.
61. See Section C of the online appendix for full details.

https://doi.org/10.1017/S002081831400023X Published online by Cambridge University Press

more countries to the negotiations. The most patient subjects chose an average category number that was 0.32 categories higher than the least patient subjects.

Level-2 thinkers, who act as if they are particularly strategic responders, were also more likely to invite more countries compared with both Level-0 thinkers (who act randomly in the beauty-contest game) and Level-1 thinkers (who also act unstrategically, essentially treating other players as a random variable). Level 2 thinkers, by contrast, chose an average category number that was 0.28 categories higher than Level-0 subjects, and 0.21 categories higher than Level-1 subjects.

We also find that women in this sample (59 percent of the university population and 21 percent of the elites) were more inclined to prefer negotiations in large numbers, though we have no explanation for why.

The Decision to Join a Treaty

Subjects' willingness to join a negotiated trade agreement depended on an enforcement mechanism as well as subjects' patience and strategic reasoning. The first model in Table 3 reports the effect of enforcement by itself. Subjects randomly assigned to the treatment in which the treaty included an enforcement mechanism were about 5 percent more likely to prefer to join the treaty.[62] The presence of an enforcement mechanism is a boon to cooperation in this sample of decision makers.

The second model reports that patience had no appreciable effect on subjects' propensity to join. This result, contrary to our hypothesis, may reflect that joining decisions arise at a single moment in time for which the treaty's attributes (its costs and benefits as well as its design, such as enforcement) play a much larger role in determining whether individuals favor membership. It may also reflect that although the question specified that some benefits would be realized only later in time it did not specify a time delay for all of the treaty's benefits. Thus, while we would expect patience to play a role, the effect may have been diluted by other considerations.

The third model reports that more strategic subjects (those measured to be Level-2 reasoners) were, on average, 11 percent more likely to join the trade agreement. As we report in the supplementary information,[63] subjects' comments reflected the intuition that strategic reasoners are more cooperative in iterated assurance games, with more strategic reasoners arguing that joining would make it more desirable for other countries to do the same. This effect is about two times the effect of adding enforcement to a trade agreement, and thus suggests that behavioral traits such as strategic thinking can exert a substantial influence on decision makers' preferences for

62. Technically subjects picked one out of five categories for how likely they were to sign the treaty. Each category was stated to represent a twenty-point probability interval (0 to 20 percent, 21 to 40 percent, etc.). Our regression assumes that movement across the five categories is linear. So 0.244 represents $20 \times 0.244 \approx 5$ percent.

63. See online appendix.

https://doi.org/10.1017/S002081831400023X Published online by Cambridge University Press

cooperation relative to the presence of enforcement deemed important by much of the literature. We cannot guarantee that this relative effect size would hold outside the lab. A host of factors may change it. However, we interpret this effect as evidence that behavioral traits can have a relatively large impact on such preferences relative to variables more commonly considered by international relations scholars. It provides a baseline for future research, which might consider how other factors make behavioral traits more or less relevant to actual elite decision making.

TABLE 3. *Decision to join trade agreement*

	Dependent variable: Would subject join treaty?			
	Enforcement	*Patience*	*Level-K*	*Interactions + controls*
ENFORCEMENT	0.244***	0.240**	0.239**	0.402**
	(0.089)	(0.090)	(0.087)	(0.164)
PATIENCE		0.008		0.002
		(0.008)		(0.013)
LEVEL-1 REASONER			0.048	0.190
			(0.092)	(0.150)
LEVEL-2 REASONER			0.546***	0.610**
			(0.147)	(0.247)
FEMALE				−0.016
				(0.101)
INCOME				0.003
				(0.021)
YEAR IN SCHOOL				0.025
				(0.031)
GRADUATE STUDENT				0.268*
				(0.157)
ENFORCEMENT * PATIENCE				0.003
				(0.018)
ENFORCEMENT * LEVEL 1				−0.340
				(0.208)
ENFORCEMENT * LEVEL 2				−0.254
				(0.323)
Intercept	3.141***	3.089***	3.077***	2.892***
	(0.068)	(0.088)	(0.074)	(0.171)
R^2	0.01	0.02	0.04	0.05
N	507	507	507	485[1]

Notes: Standard errors in parentheses are White's robust standard errors. *** $p < .01$; ** $p < .05$; * $p < .1$.
1. This regression has fewer observations because some subjects did not report their family's income.

The fourth model in Table 3 shows that these findings are robust to the addition of demographic controls. It also shows that behavioral traits do not reliably interact with the presence or absence of an enforcement mechanism. That is, the behavioral traits that we measure have an effect on the preference to join that is largely independent of whether the agreement includes a credible enforcement mechanism.

In addition to collecting quantitative evidence, we also asked subjects to comment on their decisions and reasoning processes.[64] Subjects who were classified as

64. We reproduce some of those comments in Section D of the online appendix.

https://doi.org/10.1017/S002081831400023X Published online by Cambridge University Press

nonstrategic (measured as a Level-1 reasoner) typically focused on how their country individually benefited from the treaty's provisions, or acted as though other countries' decisions were exogenous. Meanwhile, strategic (Level-2) reasoners were more likely to comment on how their decision might influence the decisions of other countries.

Elite Decision Makers

Because laboratory experiments in political science, psychology, and economics often use convenience samples of university students,[65] there is always a question of external validity, especially on matters such as voting on Congress or negotiating treaties that are exclusively the domain of elite decision makers who are difficult to engage in survey and experimental research.

To our knowledge, there are no experimental studies of elite political decision makers that measure the behavioral traits of patience; moreover, the literature on elite strategic behavior is only suggestive.[66] One study has surveyed the existing literature comparing university students with people in the general population across a number of laboratory games.[67] That study finds that differences tend to be minor and quantitative, not qualitative. That is, the same player traits exist in each population, and are simply distributed differently. The few studies that have focused on behavioral traits in depth find similar results.[68] Nonetheless, it is possible that there will be some differences. For example, there may be selection effects that influence which individuals ultimately occupy elite roles. Careers in international law and business typically have long trajectories, and patient people who highly value future success may be disproportionately represented in these careers. And, we might hope, elites who have become leaders charged with the responsibility of brokering and managing international cooperation should be more advanced strategic thinkers.[69] Similarly, elites gain practical experience on the job, and that experience

65. For examples of studies that rely on university student populations and seek to make inferences about the behavior of elites, see Ensley, de Marchi, and Munger 2007; Tingley and Walter 2011a and 2011b; and Tingley 2011.

66. To our knowledge, only three studies have looked at this—none directly focused on elite decision makers relevant for international relations. Camerer 2003, 217, citing an unpublished 1998 Camerer manuscript, looks at how undergraduates, trustees at a leading university, and CEOs score on Level-K studies. See also Plott 1996; and Bosch-Domènech et al. 2002, 1694.

67. Belot, Duch, and Miller 2010.

68. For example, Camerer, Ho, and Chong's 2003 data on the beauty contest game played among different sets of individuals (from highly experienced traders to economic PhD students to CEOs) show some differences as well as stable patterns. Even though economic PhD students play strategies that are closer to the equilibrium strategy than general student populations, they also play nonequilibrium strategies, and the pattern of nonequilibrium strategies played is similar to undergrads.

69. There is some suggestive evidence that political systems with high levels of accountability—democracies—tend to select leaders with more advanced training (and thus presumably higher levels of patience and possibly greater strategic skills). See Besley and Reynal-Querol 2011; but see also Galasso and Nannicini 2011; and Besley 2005. There is also a growing body of work that looks at characteristics of

https://doi.org/10.1017/S002081831400023X Published online by Cambridge University Press

might affect their choice of heuristics and other mechanisms for decision making even if they do not affect underlying behavioral traits. Through such mechanisms there may arise differences in the distribution of behavioral traits and related decision-making skills across elite and nonelite populations.

To evaluate whether actual policy elites with experience in the field of international trade cooperation (such as the people who have actually negotiated major trade agreements) are differently affected by our treatment or the traits we have measured, we have developed a unique sample of high-ranking government and business elites (n = 92), including former members of US Congress, their senior staff, top US trade negotiators, senior executives in firms whose operations are exposed to international trade, and civil servants in federal and state government. We built the sample through professional contacts and then expanded the sample with personal introductions. The overall response rate was almost 20 percent.[70] In the online appendix for this article, we report demographic summaries for the sample. Because it was not possible to survey these elites in our lab, we had elite subjects participate over the Web using the Qualtrics platform between April 2011 and January 2013. They participated in the identical time-discounting task and beauty-contest game and answered the identical question about whether or not they wanted to join a treaty. The elite sample is made up of individuals with an average age of fifty-two who have an average of twenty-one years' experience in business or government strategy. Like our university sample, this population is a convenience sample recruited through professional networks. Therefore, we cannot say that this sample's measured traits or question responses are representative of all professionals in the area of international trade and negotiation. However, we can examine how sensitive our findings are to whether individuals come from a convenience sample of university students or from a convenience sample of individuals with directly relevant experience. Our elite sample is necessarily much smaller than our university sample because these professionals are much busier than university students, as well as harder to contact and convince. It is therefore more difficult to recruit these individuals in large numbers.

The measured patience and Level-K reasoning of elites in our sample appear in Figures 4 and 5. As one might expect, elites differ from university students in their distribution of patience and strategic reasoning. Figure 4 shows an elite distribution that is, on average, more patient than university students. This difference is statistically significant (Wilcoxon rank sum test p-value = 1.8×10^{-6}). Figure 5 shows that elites in our sample also display a higher level of strategic reasoning, with more individuals categorized as Levels 1 and 2, and many fewer categorized as Level-0 reasoners (Wilcoxon rank sum test p-value = 4.94×10^{-6}). We do not

agents, such as international judges, in international relations and whether they are selected strategically by principles. See Voeten 2008.

70. As an experiment, during our sample creation process we also sent identical invitations to a sample of Congressional staff working on these topics but who were not familiar with our lab. The response rate was 0 percent, underscoring that building elite samples will require social network strategies.

https://doi.org/10.1017/S002081831400023X Published online by Cambridge University Press

attempt to explain here why elites differ from university students—although elsewhere we have explored what little is known about this question through a survey of the existing research in cognitive psychology and behavioral economics.[71] However, despite these differences there is still substantial overlap in both distributions. Furthermore, differences in the distribution of traits do not substantially change the relationship between those traits and the preferences that individuals have for treaties.

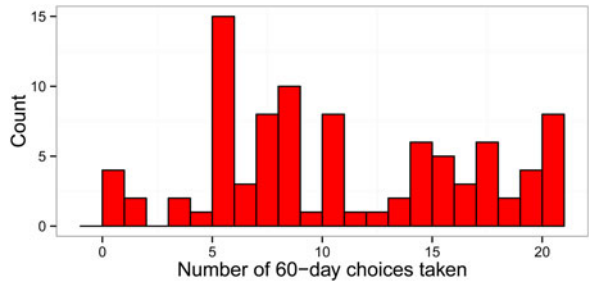


FIGURE 4. *Distribution of patience in the elite sample*

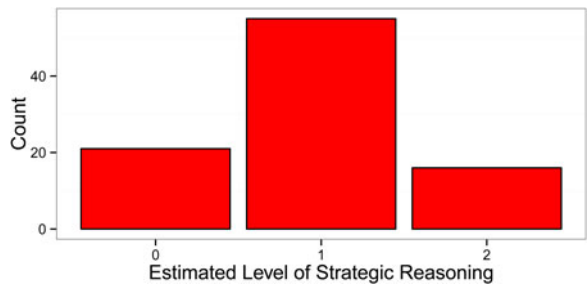


FIGURE 5. *Distribution of Level-K reasoning in the elite sample*

Table 4 displays results similar to those of Table 2, pooling both university student and elite responses to our question of whether or not to join a trade agreement. In the first three columns we look for whether there is an interaction between an individual's elite status and our treatment (column 1), their level of patience (column 2), or their Level-K reasoning (column 3). Column 4 includes all three interaction terms. Again, the results in Table 4 point in the same direction as the university-only sample in Table 2. Patience does not affect an individual's decision to join the treaty, and

71. Hafner-Burton, Hughes, and Victor 2013.

https://doi.org/10.1017/S002081831400023X Published online by Cambridge University Press

there is not a significant interaction between patience and being an elite. Likewise, Level-2 reasoners are still more likely to join the treaty, and there is not a significant interaction between this effect and elite status.

TABLE 4. *Decision to join trade agreement pooling elite and university student samples*

	Dependent variable: Would subject join treaty?			
	Enforcement	*Patience*	*Level-K*	*All interactions*
ENFORCEMENT	0.244**	0.215**	0.220**	0.236**
	(0.090)	(0.082)	(0.082)	(0.089)
PATIENCE		0.008		0.005
		(0.008)		(0.009)
LEVEL-1 REASONER			0.047	0.040
			(0.096)	(0.095)
LEVEL-2 REASONER			0.547***	0.534**
			(0.153)	(0.154)
ELITE	0.511***	0.413*	0.322	0.419
	(0.125)	(0.227)	(0.216)	(0.298)
ELITE × ENFORCEMENT	−0.161			−0.123
	(0.215)			(0.219)
ELITE × PATIENCE		−0.002		−0.005
		(0.022)		(0.021)
ELITE × LEVEL 1			0.121	0.139
			(0.268)	(0.254)
ELITE × LEVEL 2			−0.149	−0.153
			(0.301)	(0.312)
INTERCEPT	3.141***	3.102***	3.087***	3.050***
	(0.066)	(0.084)	(0.071)	(0.090)
R^2	0.04	0.04	0.06	0.06
N	599	599	599	599

Notes: Standard errors in parentheses are White's robust standard errors. *** $p < .01$; ** $p < .05$; * $p < .1$.

Thus, at least within our sample, we cannot reject the hypothesis that traits affect elites' preferences for treaties in the same manner they affect university students.[72] Moreover, the effect of the enforcement treatment is still positive. The elites in our sample clearly did not view the presence of an enforcement mechanism as a deterrent to joining the treaty. To the contrary, enforcement significantly changed their views, making formal legal cooperation more desirable.

Discussion

Together, these findings suggest that (1) our sample of elites is on average more patient and strategic than our sample of university students, and (2) that the

72. Demographics such as age and gender do not affect these results.

https://doi.org/10.1017/S002081831400023X Published online by Cambridge University Press

relationship between both traits and both pools of subjects' preferences for international trade cooperation is relatively consistent across samples. Even though elites and university students differ, patience and strategic reasoning do not have a substantially different effect on elites' preferences to join the treaty. An analogy from American politics helps to illustrate the point: consider the well-known relationship between age and voter turnout.[73] Elites will be older than university students, but the relationship between age and turnout is the same for both groups—older elites are more likely to vote than younger elites, just as older university students are more likely to vote than younger students. Our findings suggest that elites, who are more strategic and patient, prefer to negotiate in greater numbers and to join treaties more than university students do. In other words, real political elites are more favorable to large-scale cooperation.

Our findings have important implications concerning theoretical assumptions surrounding individual strategic behavior. First, the fact that elites, unlike university students, are predominantly Level-1 players may be quite important for theories in international relations. Level-1 players are strategic in the sense that they best respond to both the rules of the game and to what other players could potentially do. International relations theories that apply to political decision makers of the type in our sample may be on firmest ground when they assume that decision makers purposefully and carefully make choices with respect to what other decision makers might do, avoiding strategies that are strictly dominated. However, either because of limited cognition or strategic uncertainty, most political elites we sampled do not behave like equilibrium thinkers who act according to common knowledge rationality. Thus, theories that rely on the assumption that these kinds of elites will reliably reason[74] their way toward a particular equilibrium of a game may benefit from checking whether their results would differ if decision makers acted as Level-1 players. This is especially important when an interaction occurs relatively infrequently, such that decision makers are less likely to learn their way toward a particular equilibrium. As an example, formal models of Level-K behavior have found that Level-1 actors are more likely to be influenced by cheap-talk signals,[75] and such limited strategic reasoning could help explain puzzling phenomena in international relations, such as why cheap talk appears to work both in the real world, as well as in controlled laboratory experiments.[76]

Second, as with previous studies,[77] our findings suggest that relying on university subjects can bias the results of studies in international relations. In the strategic choice we studied (of whether or not to join a trade agreement), more strategic individuals

73. Powell 1986.
74. This is not the same as saying that actors will not reach the equilibrium of the game through processes such as learning or evolution—it is simply that they may not always reach the equilibrium of a new game through sheer reasoning.
75. Crawford 2003.
76. Tingley and Walter 2011a.
77. For example, Mintz, Redd, and Vedlitz 2006.

https://doi.org/10.1017/S002081831400023X Published online by Cambridge University Press

are more cooperative, and relying solely on university students (who are, on average, less strategic than elites) would systematically underestimate leaders' willingness to join trade agreements. In other strategic settings where cooperation is not a strategic complement, more strategic individuals can actually have a higher threshold for acting cooperatively—in these scenarios, studies using university students could have the opposite bias, overestimating actors' propensity to cooperate.

Our findings represent a first step in more precisely identifying how policy elites are different from a university student sample, and potentially allow researchers to correct for some of the biases that can arise from studying university students as proxies for elite behavior. For example, the effect of Level-K reasoning—at least in our study—appears to be the same in both populations. This finding opens up a potential alternative to producing highly costly elite samples such as we have done in this study. Researchers could reweight or match a university student sample such that they look more like elites on theoretically relevant dimensions, such as strategic reasoning or patience. That approach would require administering games to elicit each player's Level-K type (which will add a few additional minutes onto a survey) and then use the results to calibrate findings to elite characteristics. Despite this inconvenience, such an approach could allow scholars to better generalize findings from students or other nonelite samples (which are much easier to obtain and less costly to survey) to the policy elites who predominately make decisions in international relations.

Preferences to Political Outcomes

This article is about how different behavioral traits shape the preferences of individuals, not the ultimate outcomes of international cooperation such as the level of international trade or actual trade cooperation. Thus, it is a first step in untangling how behavioral traits could shape international relations decisions. For many foreign policy decisions, however, decision makers work in teams and are embedded in bureaucracies, both of which can mute the causal link between any single person's preferences for cooperation and real-world policy outcomes. Nonetheless, individual preferences for cooperation and their determinants are important to study for at least four reasons, and may help to shed light on some actual political decisions that affect international relations.

First, scholars long ago noted that the people sitting at the negotiation table and those who make decisions to join agreements can have substantial autonomy on their own—they are not merely perfect agents for underlying interests and structures.[78] They have personal styles, opinions, and predilections, as well as formal and informal permissions and job flexibility, which give them degrees of independence from their principals.[79] This agency slack creates room for behavioral traits to matter, and—where slack is present—makes it unlikely that governmental

78. See Putnam 1988; Evans, Jacobson, and Putnam 1993; and Shamir and Shikaki 2005.
79. Hawkins et al. 2006.

https://doi.org/10.1017/S002081831400023X Published online by Cambridge University Press

decisions are completely determined by external circumstances, such that any set of agents would do the same thing in the same situation.

Second, within collective decision-making processes more generally, there is already evidence that individuals can sometimes have a large effect on agendas and decision-making procedures.[80] For instance, when the talks that eventually led to the creation of the WTO dragged past their deadline and were stuck in gridlock, the chief of the negotiating process, Arthur Dunkel, took control by creating a single negotiating text (the Dunkel Draft) that, with very few revisions, countries finally accepted.

Third, the importance of individuals in group reasoning depends on the kind of reasoning involved. When the decisions to be made have clear evaluation criteria —for example, equilibrium reasoning—groups behave systematically more rationally than individuals.[81] For those kinds of tasks, group outcomes may not be particularly sensitive to the exact composition of the group because members of the group that do not grasp the concepts are likely to follow the reasoning of group members who do. In contrast, when group decision making requires judgmental tasks where outcomes hinge on weighing uncertainty, human values, perceptions of strategic opponents, and other criteria for which there are no obvious correct answers then groups do not demonstrate appreciable gains compared with individuals. In these cases, the behavioral traits of the group members can have a large effect on outcomes.[82] Many of the grand decisions in international relations are of this second, judgmental, type. For example, policy-makers in the United States and several other countries are currently focused on creating small trade pacts—a decision based on the judgment that such agreements allow for progress in opening trade and will not cause undue harm by discriminating against the WTO and efforts to strengthen global trade institutions. That judgment is based on assessments of how many highly uncertain factors will unfold in the future rather than the result of any iron-clad laws about how trade policy will exactly unfold. Although more work is needed to understand how individual personality affects collective decision making on tasks relevant to international relations, the existing research preliminarily suggests that the characteristics of the people in the room—even when decisions are made in groups by established bureaucracies—can have a large impact on the outcome of committee decisions.

Fourth, key decisions in the design and implementation of international agreements often take place in small groups where a few decision makers have inordinate influence. In crafting a round of trade talks, a small group of decision makers—usually drawn from the most powerful countries and the secretariat of the WTO—make the final decision about topics to place on the agenda. The successful resolution of trade talks usually hinges on the efforts of a few negotiators—often the Director General of the WTO himself—to craft a final negotiating text from many different competing

80. See, for example, Forsyth 2010; and Bonner and Baumann 2008.
81. See Hastie 1986; Levine and Moreland 2006; and Laughlin 2011.
82. See Robert and Carnevale 1997; and Elbittar, Gomberg, and Sour 2011. For reviews, see Kugler, Kausel, and Kocher 2012; and Charness and Sutter 2012.

https://doi.org/10.1017/S002081831400023X Published online by Cambridge University Press

proposals. That text reflects political and perhaps also behavioral decisions about which elements of a trade agreement are most and least important. And decisions about ratification in important countries such as the United States require a deal that is crafted with the authority of a few leaders in the executive and legislative branches. Such decisions often take place in the context of a crisis or with a looming deadline that concentrates minds and, for efficiency, shrinks the number of people in the room.[83] This pattern is hardly unique to trade. In most other areas of complex international cooperation the final deals reached are the work of a few people—for example, brokering the main agreement at the 2009 Copenhagen Conference were the heads of state from just five countries with few other individuals sitting around the table.[84] To be sure, decision-making institutions such as bureaucracies matter for international affairs, but at pivotal moments perhaps individuals also matter.

Although understanding the determinants of decision makers' heterogeneous preferences for international cooperation may help to shed light on variations in some actual political decisions, there may also be circumstances where we should expect decision makers to produce highly similar policy choices despite very different personal dispositions and skills. For example, where policy decisions are highly constrained by circumstances—such as when implementing agreements that have already been crafted in detail or managing a crisis that affords little room to maneuver—institutions and context may matter a lot more than whether decision makers with widely varied backgrounds and personality traits prefer different outcomes. This may help explain, for example, the (highly controversial) musings that presidents as different as George W. Bush and Barack Obama are actually quite similar in important matters of economic policy (for example, managing weak banks) and international security (for example, escalation in Afghanistan).[85] Our claim is not that decision makers' behavioral traits always drive their foreign policy decisions but rather that traits may in some circumstances have a role to play in shaping how people reason through choices about international cooperation.

Conclusion

Facing the same situation and placed in the same decision-making role, our subjects' behavioral traits reliably correspond to differences in their views on international cooperation. We conclude by looking at three implications for the study of international cooperation.

83. These points are regularly emphasized by insider accounts of international negotiations. For example, on the 1992 Rio Conference, see Brenton 1994; on the US-Canada free trade talks that were precursors to the North American Free Trade Agreement (NAFTA), see Hart 1994.

84. For example, John M. Broder, "Many Goals Remain Unmet in Five Nations' Climate Deal," *New York Times*, 19 December 2009, A1.

85. David Bromwich, "Symptoms of the Bush-Obama Presidency," *Huffington Post*, 18 August 2011. Available at www.huffingtonpost.com/david-bromwich/symptoms-of-the-bushobama_b_930260.html, accessed 12 February 2014.

https://doi.org/10.1017/S002081831400023X Published online by Cambridge University Press

First, existing research in international relations has reflected on what happens if states are more or less patient or strategic. Our study complements this literature by suggesting that states may depend on the patience and strategic reasoning of the individuals making the key decisions about cooperation. Research on the micro-level sources of decision makers' preferences would help identify why cooperation processes are, at times, far from perfectly optimal or functional to the strategic problem at hand. The evidence from our study suggests that these behavioral traits could have an influence on decision makers' preferences that is equal to or larger than other factors that have attracted intense scholarly attention, such as enforcement of agreements.

Second, work of this type might help shed light on principal-agent (P-A) relationships in international relations. In the tradition of P-A frameworks, it is assumed that when there is slack between the underlying principal and the agents the agents pursue their own interests—seeking to gain privately at the expense of serving the principal's goals.[86] Although that behavior may be important in some settings, the work here suggests agency slack might operate in other ways as well. At critical junctions in negotiations—when there are a few agents in the room and where the outcomes reflect in large measure their judgment—these factors could prove to be very important and also amenable to systematic study.

Third, this work suggests an agenda for research that links the revolution in methods and insights from behavioral economics and cognitive psychology to the study of international relations, including institutions. A full description of that agenda is beyond this article, but one promising direction suggested by our research is the potential for using convenience samples to study at least some of these phenomena. Our article suggests that the relationship between strategic reasoning and joining a cooperative trade agreement is robust to whether one uses a convenience sample of undergraduates or elites. It is quite plausible that this is not the only area in international relations where decision-making traits are distributed differently, but do not function differently. When this relationship holds, scholars of international relations could sample from the part of a student population that most closely approximates the traits of actual policy-makers, leading to research that is more likely to have external validity.

Supplementary material

Replication data are available at http://dx.doi.org/10.1017/S002081831400023X.

References

Ackerman, Phillip L., and Walter Schneider. 1985. Individual Differences in Automatic and Controlled Information Processing. In *Individual Differences in Cognition*. Vol. 2, edited by Ronna F. Dillon, 35–66. New York: Academic Press.

86. See Lake and McCubbins 2006.

https://doi.org/10.1017/S002081831400023X Published online by Cambridge University Press

Agranov, Marina, Andrew Caplin, and Chloe Tergiman. 2013. Naïve Play and the Process of Choice in Guessing Games. Mimeo. Pasadena: California Institute of Technology and the Center for Experimental Social Science.

Agranov, Marina, Elizabeth Potamites, Andrew Schotter, and Chloe Tergiman. 2012. Beliefs and Endogenous Cognitive Levels: An Experimental Study. *Games and Economic Behavior* 75 (2):449–63.

Alter, Karen J., and Sophie Meunier. 2009. The Politics of International Regime Complexity. *Perspectives on Politics* 7 (1):13–24.

Axelrod, Robert, and Robert Keohane. 1985. Achieving Cooperation Under Anarchy: Strategies and Institutions. *World Politics* 38 (1):226–54.

Belot, Michèle, Raymond Duch, and Luis Miller. 2010. Who Should Be Called to the Lab? A Comprehensive Comparison of Students and Non-Students in Classic Experimental Games. Centre for Experimental Social Sciences Discussion Paper. Oxford, UK: Nuffield College, University of Oxford.

Berinsky, Adam J. 2004. Can We Talk? Self-Presentation and the Survey Response. *Political Psychology* 25 (4):643–59.

Besley, Timothy. 2005. Political Selection. *Journal of Economic Perspectives* 19 (3):43–60.

Besley, Timothy, and Marta Reynal-Querol. 2011. Do Democracies Select More Educated Leaders? *American Political Science Review* 105 (3):552–66.

Bhui, Rahul, and Colin F. Camerer. 2011. Measuring Intrapersonal Stability of Strategic Sophistication in Cognitive Hierarchy Modeling. Mimeo. Pasadena: California Institute of Technology.

Boli, John, and George M. Thomas. 1999. *Constructing World Culture: International Nongovernmental Organizations Since 1875*. Stanford, CA.: Stanford University Press.

Bonner, Bryan L., and Michael R. Baumann. 2008. Informational Intra-Group Influence: The Effects of Time Pressure and Group Size. *European Journal of Social Psychology* 38 (1):44–66.

Bosch-Domènech, Antoni, Jose G. Montalvo, Rosemarie Nagel, and Albert Satorra. 2002. One, Two, (Three), Infinity …: Newspaper and Lab Beauty-Contest Experiments. *American Economic Review* 92 (5):1687–701.

Brenton, Tony. 1994. *The Greening of Machiavelli: The Evolution of International Environmental Politics*. London: Royal Institute of International Affairs, Energy and Environmental Programme.

Byman, Daniel L., and Kenneth M. Pollack. 2001. Let Us Now Praise Great Men: Bringing the Statesman Back in. *International Security* 25 (4):107–46.

Camerer, Colin F. 2003. *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton, NJ: Princeton University Press.

Camerer, Colin F., and Robin M. Hogarth. 1999. The Effects of Financial Incentives in Experiments: A Review and Capital-Labor-Production Framework. *Journal of Risk and Uncertainty* 19 (1–3):7–42.

Camerer, Colin, Teck-Hua Ho, and Kuan Chong. 2003. Models of Thinking, Learning, and Teaching in Games. *American Economic Review* 93 (2):192–95.

———. 2004. A Cognitive Hierarchy Model of Games. *Quarterly Journal of Economics* 119 (3):861–98.

Card, Andrew H., Thomas A. Daschle, Edward H. Alden, and Matthew J. Slaughter. 2011. *US Trade and Investment Policy*. Independent Task Force Report No. 67. Council on Foreign Relations.

Charness, Gary, and Matthias Sutter. 2012. Groups Make Better Self-Interested Decisions. *Journal of Economic Perspectives* 26 (3):157–76.

Chen, Serena, and Shelly Chaiken. 1999. The Heuristic-Systematic Model in Its Broader Context. In *Dual-Process Theories in Social Psychology*, edited by Shelly Chaiken and Yaacov Trope, 73–96. New York: Guilford Press.

Coller, Maribeth, and Melonie B. Williams. 1999. Eliciting Individual Discount Rates. *Experiment Economics* 2 (2):107–27.

Cooper, David J., John H. Kagel, Wei Lo, and Qing Liang Gu. 1999. Gaming Against Managers in Incentive Systems: Experimental Results with Chinese Students and Chinese Managers. *American Economic Review* 89 (4):781–804.

Coricelli, Giorgio, and Rosemarie Nagel. 2009. Neural Correlates of Depth of Strategic Reasoning in Medial Prefrontal Cortex. *Proceedings of the National Academy of Sciences* 106 (23):9163–38.

https://doi.org/10.1017/S002081831400023X Published online by Cambridge University Press

Costa-Gomes, Miguel A., and Vincent P. Crawford. 2006. Cognition and Behavior in Two-Person Guessing Games: An Experimental Study. *American Economic Review* 96 (5):1737–68.
Costa-Gomes, Miguel A., and Klaus G. Zauner. 2001. Ultimatum Bargaining Behavior in Israel, Japan, Slovenia, and the United States: A Social Utility Analysis. *Games and Economic Behavior* 34 (2): 238–69.
Crawford, Vincent P. 2003. Lying for Strategic Advantage: Rational and Boundedly Rational Misrepresentation of Intentions. *American Economic Review* 93 (1):133–49.
Crawford, Vincent P., Miguel A. Costa-Gomes, and Nagore Iriberri. 2013. Structural Models of Nonequilibrium Strategic Thinking: Theory, Evidence, and Applications. *Journal of Economic Literature* 51 (1):5–62.
Dawes, Christopher T., Peter J. Loewen, and James H. Fowler. 2011. Social Preferences and Political Participation. *Journal of Politics* 73 (3):845–56.
De Sousa, José, Guillaume Hallard, and Antoine Terracol. 2012. Do Non-strategic Players Really Exist? Evidence From Experimental Games Involving Step Reasoning. Mimeo. Paris: Université de Paris Sud and Paris School of Economics.
Donno, Daniela. 2010. Who Is Punished? Regional Intergovernmental Organizations and the Enforcement of Democratic Norms. *International Organization* 64 (4):593–625.
Downs, George W., David M. Rocke, and Peter N. Barsoom. 1996. Is the Good News About Compliance Good News About Cooperation? *International Organization* 50 (3):379–406.
Elbittar, Alexander, Andrei Gomberg, and Laura Sour. 2011. Group Decision-Making and Voting in Ultimatum Bargaining: An Experimental Study. *B.E. Journal of Economic Analysis and Policy* 11 (1): Article 53.
Ensley, Michael, Scott de Marchi, and Michael Munger. 2007. Candidate Uncertainty, Mental Models, and Complexity: Some Experimental Results. *Public Choice* 132 (1):231–46.
Evans, Peter B., Harold K. Jacobson, and Robert D. Putnam, eds. 1993. *Double-Edged Diplomacy: International Bargaining and Domestic Politics*. Berkeley: University of California Press.
Fearon, James. 1998. Bargaining, Enforcement, and International Cooperation. *International Organization* 52 (2):269–305.
Fehr, Ernst, and John A. List. 2004. The Hidden Costs and Returns of Incentives—Trust and Trustworthiness Among CEOs. *Journal of the European Economic Association* 2 (5):743–71.
Forsyth, Donelson R. 2010. *Group Dynamics*. 5th ed. Belmont, CA: Wadsworth.
Fowler, James H. 2006. Altruism and Turnout. *Journal of Politics* 68 (3):674–83.
Fowler, James H., and Cindy D. Kam. 2006. Patience as a Political Virtue: Delayed Gratification and Turnout. *Political Behavior* 28 (2):113–28.
———. 2007. Beyond the Self: Altruism, Social Identity, and Political Participation. *Journal of Politics* 69 (3):813–27.
Fowler, James H., and Darren Schreiber. 2008. Biology, Politics, and the Emerging Science of Human Nature. *Science* 322 (5903):912–14.
Galasso, Vincenzo, and Tommaso Nannicini. 2011. Competing on Good Politicians. *American Political Science Review* 105 (1):79–99.
Gallup Organization, and Times Mirror Company. 1989. The People, the Press and Politics: Public Opinion About Economic Issues: Viewpoints of the General Public, US Opinion Leaders, and Overseas Investors. Available at <http://www.people-press.org/files/legacy-pdf/19890301.pdf>. Accessed 14 February 2014.
Gartner, Scott S. 2011. On Behalf of a Grateful Nation: Conventionalized Images of Loss and Individual Opinion Change in War. *International Studies Quarterly* 55 (2):545–61.
Gerber, Alan, Gregory Huber, David Doherty, Conor M. Dowling, and Shang E. Ha. 2010. Personality and Political Attitudes: Relationships Across Issue Domains and Political Contexts. *American Political Science Review* 104 (1):111–33.
Gilligan, Michael J. 2006. Is Enforcement Necessary for Effectiveness? A Model of the International Criminal Regime. *International Organization* 60 (4):935–67.

https://doi.org/10.1017/S002081831400023X Published online by Cambridge University Press

Gilligan, Michael J., and Leslie Johns. 2012. Formal Models of International Institutions. *Annual Review of Political Science* 15 (1):221–43.
Greenstein, Fred I. 2004. *The Presidential Difference: Leadership Style from FDR to George W. Bush*. Princeton, NJ: Princeton University Press.
Grieco, Joseph M., Christopher Gelpi, Jason Reifler, and Peter D. Feaver. 2011. Let's Get a Second Opinion: International Institutions and American Public Support for War. *International Studies Quarterly* 55 (2):563–83.
Hafner-Burton, Emilie M., D. Alex Hughes, and David G. Victor. 2013. The Cognitive Revolution and the Political Psychology of Elite Decision Making. *Perspectives on Politics* 11 (2):368–86.
Hafner-Burton, Emilie M., David G. Victor, and Yonatan Lupu. 2012. Political Science Research on International Law: The State of the Field. *American Journal of International Law* 106 (1):47–97.
Hainmueller, Jens, and Michael J. Hiscox. 2010. Attitudes Toward Highly Skilled and Low-Skilled Immigration: Evidence from a Survey Experiment. *American Political Science Review* 104 (1):61–84.
Harrison, Glenn W., Morton I. Lau, and Melonie B. Williams. 2002. Estimating Individual Discount Rates in Denmark: A Field Experiment. *American Economic Review* 92 (5):1606–17.
Hart, Michael. 1994. *Decision at Midnight: Inside the Canada-US Free-Trade Negotiations*. Vancouver, BC: University of British Columbia Press.
Hastie, Reid. 1986. Experimental Evidence on Group Accuracy. In *Information Pooling and Group Decision Making: Proceedings of the Second University of California, Irvine, Conference on Political Economy*, edited by Bernard Grofman and Guillermo Owen, 129–57. Greenwich, CT: JAI Press.
Hatemi, Peter K., and Rose McDermott. 2012. A Neurobiological Approach to Foreign Policy Analysis: Identifying Individual Differences in Political Violence. *Foreign Policy Analysis* 8 (2):111–29.
Hawkins, Darren G., David A. Lake, Daniel L. Nielson, and Michael J. Tierney, eds. 2006. *Delegation and Agency in International Organizations*. Cambridge, MA: Cambridge University Press.
Hedinger, Beat, and Lorenz Götte. 2006. Cooperation in the Cockpit: Evidence of Reciprocity and Trust Among Swiss Air Force Pilots. Unpublished manuscript, Institute for Strategy and Business Economics, University of Zurich, Switzerland.
Helland, Leif, Gjermund Loyning, and Lars C. Monkerud. 2013. Don't You Believe Me? Seasoned Parliamentarians in a Lobbying Experiment. Working Paper. Oslo: Norwegian Business School.
Hills, Carla A. 2005. The Stakes of Doha. *Foreign Affairs* (Special WTO ed.) 84 (7):25–36.
Hug, Simon, and Thomas König. 2002. In View of Ratification: Governmental Preferences and Domestic Constraints at the Amsterdam Intergovernmental Conference. *International Organization* 56 (2):447–76.
Jervis, Robert. 1968. Hypotheses on Misperception. *World Politics* 20 (3):454–79.
———. 1976. *Perception and Misperception in International Politics*. Princeton, NJ: Princeton University Press.
———. 1998. *System Effects: Complexity in Political and Social Life*. Princeton, NJ: Princeton University Press.
Johnson, Dominic D.P. 2004. *Overconfidence and War: The Havoc and Glory of Positive Illusions*. Cambridge, MA: Harvard University Press.
Kagan, Jerome. 1988. The Meanings of Personality Predicates. *American Psychologist* 43 (8):614–20.
Kahler, Miles. 1992. Multilateralism with Small and Large Numbers. *International Organization* 46 (3): 681–708.
Keohane, Robert O. 2005. *After Hegemony: Cooperation and Discord in the World Political Economy*. Princeton, NJ: Princeton University Press.
Keohane, Robert O., and David G. Victor. 2011. The Regime Complex for Climate Change. *Perspective on Politics* 9 (1):7–23.
Knetsch, Jack L. 1989. The Endowment Effect and Evidence of Nonreversible Indifference Curves. *American Economic Review* 79 (5):1277–84.
Kugler, Tamar, Edgar E. Kausel, and Martin G. Kocher. 2012. Are Groups More Rational Than Individuals? A Review of Interactive Decision Making in Groups. CESifo Working Paper Series 3701. Munich, Germany: CESifo Group Munich.

https://doi.org/10.1017/S002081831400023X Published online by Cambridge University Press

Lake, David A., and Mathew D. McCubbins. 2006. The Logic of Delegation to International Organizations. In *Delegation and Agency in International Organizations*, edited by Darren G. Hawkins, David A. Lake, Daniel L. Nielson, and Michael J. Tierney, 341–68. Cambridge, UK: Cambridge University Press.

Laughlin, Patrick R. 2011. *Group Problem Solving*. Princeton, NJ: Princeton University Press.

Le Coq, Chloé, and Jon Thor Sturluson. 2012. Does Opponents' Experience Matter? Experimental Evidence from a Quantity Precommitment Game. *Journal of Economic Behavior and Organization* 84 (1):265–77.

Leventoğlu, Bahar, and Ahmer Tarar. 2008. Does Private Information Lead to Delay or War in Crisis Bargaining? *International Studies Quarterly* 52 (3):533–53.

Levine, John M., and Richard L. Moreland. 2006. *Small Groups: Key Readings*. 1st ed. New York: Psychology Press.

List, John A. 2003. Does Market Experience Eliminate Market Anomalies? *Quarterly Journal of Economics* 118 (1):41–71.

List, John A. and Charles F. Mason. 2009. Are CEOs Expected Utility Maximizers? *Journal of Econometrics* 162 (1):114–23.

Mansfield, Edward D., and Jon C. Pevehouse. 2008. Democratization and the Varieties of International Organizations. *Journal of Conflict Resolution* 52 (2):269–94.

Martin, Lisa L. 1995. Heterogeneity, Linkage and Commons Problems. In *Local Commons and Global Interdependence*, edited by Robert O. Keohane and Elinor Ostrom, 79–91. London: Sage.

McDermott, Rose. 2008. *Presidential Leadership, Illness, and Decision Making*. New York: Cambridge University Press.

———. 2011. New Directions for Experimental Work in International Relations. *International Studies Quarterly* 55 (2):503–20.

Mearsheimer, John. 1994. The False Promise of International Institutions. *International Security* 19 (3):5–49.

Meyer, John W., John Boli, George M. Thomas, and Francisco O. Ramirez. 1997. World Society and the Nation-State. *American Journal of Sociology* 103 (1):144–81.

Milner, Helen. 1997. *Interests, Institutions, and Information: Domestic Politics and International Relations*. Princeton, NJ: Princeton University Press.

Mintz, Alex, Steven B. Redd, and Arnold Vedlitz. 2006. Can We Generalize From Student Experiments to the Real World in Political Science, Military Affairs, and International Relations? *Journal of Conflict Resolution* 50 (5):757–76.

Mintz, Alex, Yi Yang, and Rose McDermott. 2011. Experimental Approaches to International Relations. *International Studies Quarterly* 55 (2):493–501.

Mischel, Walter. 1974. Processes in Delay of Gratification. In *Advances in Experimental Social Psychology*. Vol. 7, edited by Roger Berkowitz, 249–92. New York: Academic Press.

Mondak, Jeffery J., and Karen D. Halperin. 2008. A Framework for the Study of Personality and Political Behaviour. *British Journal of Political Science* 38:335–62.

Mondak, Jeffery J., Matthew V. Hibbing, Damarys Canache, Mitchell A. Seligson, and Mary R. Anderson. 2010. Personality and Civic Engagement: An Integrative Framework for the Study of Trait Effects on Political Behavior. *American Political Science Review* 104 (1):85–110.

Myatt, David P., Hyun Song Shin, and Chris Wallace. 2002. The Assessment: Games and Coordination. *Oxford Review of Economic* Policy 18 (4):397–417.

Nagel, Rosemarie. 1995. Unraveling in Guessing Games: An Experimental Study. *American Economic Review* 85 (5):1313–26.

Neale, Margaret A., and Max H. Bazerman. 1985. The Effects of Framing and Negotiator Overconfidence on Bargaining Behaviors and Outcomes. *Academy of Management Journal* 28 (1):34–49.

Neustadt, Richard E. 1960. *Presidential Power and the Modern Presidents: The Politics of Leadership From Roosevelt to Reagan*. New York: Wiley.

Oye, Kenneth A. 1985. Explaining Cooperation Under Anarchy: Hypotheses and Strategies. *World Politics* 38 (1):1–24.

https://doi.org/10.1017/S002081831400023X Published online by Cambridge University Press

Palfrey, Thomas R., and Stephanie W. Wang. 2009. On Eliciting Beliefs in Strategic Games. *Journal of Economic Behavior and Organization* 71 (2):98–109.

Plott, Charles R. 1996. Rational Individual Behavior in Markets and Social Choice Processes: The Discovered Preference Hypothesis. In *The Rational Foundations of Economic Behavior: Proceedings of the IEA Conference Held in Turin, Italy*, edited by Kenneth Arrow, Enrico Colombatto, Mark Perlman, and Christian Schmidt, 220–24. London: Macmillan.

Powell, G. Bingham, Jr. 1986. American Voter Turnout in Comparative Perspective. *American Political Science Review* 80 (1):17–43.

Powell, Robert. 1999. *In the Shadow of Power*. Princeton, NJ: Princeton University Press.

Putnam, Robert D. 1988. Diplomacy and Domestic Politics: The Logic of Two-Level Games. *International Organization* 42 (3):427–60.

Putnam, Tonya L. and Jacob N. Shapiro. 2009. Does International Law Affect Willingness to Punish Foreign Human Rights Abusers? Working Paper. Princeton University/Columbia University, Princeton, NJ/New York.

Raustiala, Kal. 2005. Form and Substance in International Agreements. *American Journal of International Law* 99 (3):581–614.

Robert, Christopher, and Peter J. Carnevale. 1997. Group Choice in Ultimatum Bargaining. *Organizational Behavior and Human Decision Processes* 72 (2):256–79.

Sandler, Todd. 2004. *Global Collective Action*. Cambridge, UK: Cambridge University Press.

Schlesinger Jr., Arthur M. 1965. *A Thousand Days: John F. Kennedy in the White House*. New York: Houghton Mifflin.

Schwarz, Norbert. 1999. Self-Reports: How the Questions Shape the Answers. *American Psychologist* 54 (2):93–105.

Shamir, Jacob, and Khalil Shikaki. 2005. Public Opinion in the Israeli-Palestinian Two-Level Game. *Journal of Peace Research* 42 (3):311–28.

Smith, James McCall. 2000. The Politics of Dispute Settlement Design: Explaining Legalism in Regional Trade Pacts. *International Organization* 54 (1):137–80.

Stahl, Dale O., and Paul W. Wilson. 1995. On Players' Models of Other Players: Theory and Experimental Evidence. *Games and Economic Behavior* 10 (1):218–54.

Stein, Arthur A. 1982. Coordination and Collaboration: Regimes in an Anarchic World. *International Organization* 36 (2):299–324.

Tingley, Dustin H. 2011. The Dark Side of the Future: An Experimental Test of Commitment Problems in Bargaining. *International Studies Quarterly* 55:521–44.

Tingley, Dustin H., and Stephanie W. Wang. 2010. Belief Updating in Sequential Games of Two-Sided Incomplete Information: An Experimental Study of a Crisis Bargaining Model. *Quarterly Journal of Political Science* 5 (3):243–55.

Tingley, Dustin H., and Barbara Walter. 2011a. Can Cheap Talk Deter? An Experimental Analysis. *Journal of Conflict Resolution* 55 (6):994–1018.

———. 2011b. The Effect of Repeated Play on Reputation Building: An Experimental Approach. *International Organization* 65 (2):343–65.

Tomz, Michael. 2004. Interests, Information, and the Domestic Politics of International Agreements. Working Paper. Stanford, CA: Stanford University.

———. 2008. Reputation and the Effect of International Law on Preferences and Beliefs. Working Paper. Stanford, CA: Stanford University.

Trager, Robert F., and Lynn Vavreck. 2011. The Political Costs of Crisis Bargaining: Presidential Rhetoric and the Role of Party. *American Journal of Political Science* 55 (3):526–45.

United States Trade Representative (USTR). 2008. Ambassador Susan C. Schwab, USTR-Doha Media Roundtable. 17 July 2008. Audio/Transcript: USTR Schwab Doha Media Roundtable. Available at <www.ustr.gov/about-us/press-office/press-releases/archives/2008/july>. Accessed June 28 2012.

Vecchione, Michele, and Gian V. Caprara. 2009. Personality Determinants of Political Participation: The Contribution of Traits and Self-Efficacy Beliefs. *Personality and Individual Differences* 46 (4):487–92.

https://doi.org/10.1017/S002081831400023X Published online by Cambridge University Press

Voeten, Erik. 2008. The Impartiality of International Judges: Evidence from the European Court of Human Rights. *American Political Science Review* 102 (4):417–33.

Von Stein, Jana. 2005. Do Treaties Constrain or Screen? Selection Bias and Treaty Compliance. *American Political Science Review* 99 (4):611–22.

Webb, Eugene J., Donald T. Campbell, Richard D. Schwartz, and Lee Sechrest. 2000. *Unobtrusive Measures*. Thousand Oaks, CA: Sage.

Woodward, Bob. 2010. *Obama's Wars*. New York: Simon and Schuster.

Yoshida, Wako, Ray J. Dolan, and Karl J. Friston. 2008. Game Theory of Mind. *PLoS Computational Biology* 4 (12):e1000254.

Yoshida, Wako, Isabel Dziobek, Dorit Kliemann, Hauke R. Heekeren, Karl J. Friston, and Ray J. Dolan. 2010. Cooperation and Heterogeneity of the Autistic Mind. *Journal of Neuroscience* 30 (26):8815–18.

Yoshida, Wako, Ben Seymour, Karl J. Friston, and Raymond J. Dolan. 2010. Neural Mechanisms of Belief Inference During Cooperative Games. *Journal of Neuroscience* 30 (32):10744–51.

https://doi.org/10.1017/S002081831400023X Published online by Cambridge University Press