Emotions and the Micro-Foundations of Commitment Problems

Jonathan Renshon, Julia J. Lee, and Dustin Tingley

Abstract While emotions are widely regarded as integral to the "behavioral approach" to International Relations (IR), a host of fundamental problems have delayed the integration of affective influences into traditional models of IR. We aim to integrate affect by focusing on commitment problems, a body of work that contains strong theoretical predictions about how individual decision makers will and should act. Across two lab experiments, we use a novel experimental protocol that includes a psychophysiological measure of emotional arousal (skin conductance reactivity) to study how individuals react to changes in bargaining power. While we find support for one key pillar of IR theory—individuals do reject offers when they expect the opponent's power to increase—we also find that physiological arousal tampers with individuals' ability to think strategically in the manner predicted by canonical models. Our follow-up experiment mimics the elements of institutional solutions to commitment problems and finds support for their efficacy on the individual level. Our novel findings suggest that when individuals face large power shifts, emotional arousal short-circuits their ability to "think forward and induct backwards," suggesting that emotionally aroused individuals are less prone to commitment problems.

Although emotions play an "obvious and omnipresent role in world politics," several factors have delayed IR scholars from integrating affective influences into more traditional frameworks of international relations.[1] First, despite some early forays, emotions seem to have only recently been accepted as part of "mainstream" IR. In this way, the progress in our field has largely mirrored that of psychology, where theories of cognition and processing operated for many decades in isolation from theories of affect. Second has been the practical difficulty of examining emotions, leading to a proliferation of helpful theoretical pieces but little in the way of systematic evidence.[2] While researchers have caught on quickly, the new generation of emotion research in

We are especially grateful to Ariya Hagh and Anna Oltman, and the team at the Harvard Decision Science Laboratory for facilitating this research, including Judith Ezike, Dan Zangri, Nicole Ludmir, Gabe Mansur, Roseanne McManus, Jeanie Nguyen, Jakob Schneider, Phil Esterman, Claire Tan, Will Boning, Carolyn Killea, Amee Xu, Jack Schultz, Sarah Sussman, and Mark Edington. Tingley thanks Harvard's Weatherhead Center for International Affairs and Institute for Quantitative Social Science for financial support. We thank Josh Kertzer, Rose McDermott, Paul Poast, Alex Peysakhovich, Yunkyu Sohn, Jessica Weeks, and the UCSD Behavioral International Relations workshop organizers, especially David Victor, for their thoughtful comments on previous versions of this paper.

1. Bleiker and Hutchison 2008, 115.

2. This is less true for studies of the general public's emotions, where surveys and survey experiments have provided a number of insights into a wide range of topics.

International Organization 71, Supplement 2017, pp. S189–S218
© The IO Foundation, 2017.
doi:10.1017/S0020818316000473

https://doi.org/10.1017/S0020818316000473 Published online by Cambridge University Press

IR must be supplemented by studies that are more systemic in nature and that bypass some of the severe inferential problems associated with inferring emotions from documentary records and the biased recollections of leaders.

We examine emotions in the context of a foundational question that lies at the heart of the scientific study of international relations: why, given that conflict is costly and destroys resources, does it ever occur? Rationalist accounts of wars, which frequently turn on the presence of a commitment problem, are one of the most promising class of theories.[3] Commitment problems arise because shifts in power generate incentives to renege on current commitments: if A is increasing in strength relative to B, A will find it difficult to "credibly commit" to not take advantage of B in the future. With this in mind, the actor whose power is declining (B) prefers conflict while it still retains a relative power advantage rather than a peaceful resolution that may leave it open to future predation.[4]

Bargaining theory represents a useful starting place for the study of emotions since it ultimately relies on the choices of individual actors. Just as importantly, the clearly specified assumptions of bargaining theory generate strong theoretical expectations about what behaviors and choices we *should* expect, making it much easier to evaluate the influence of factors outside the model, such as emotions (or other psychological characteristics such as epistemic motivation).[5] An added advantage of our "behavioral approach"—the integration of formal models of human behavior with experimental methods focused on identifying the impact of emotions on decision making—is that it helps to improve both literatures from which we draw. Formal models are improved by subjecting their assumptions and expectations to empirical testing and learning about the impact of individual heterogeneity on outcomes, and our understanding of emotions is similarly advantaged by examining when and how they lead to outcomes not predicted by existing models.[6]

How do emotions affect the choices of individuals who face commitment problems in bargaining? Along the way to answering that question, we investigate and provide evidence that bears on two other questions, both fundamental to current theories of bargaining and commitment problems in IR. First, how does the *size of the power shift* (large or small) affect how individual decision makers respond in a bargaining situation? Second, do changes in the institutions (or rules of the game) ameliorate the dilemma posed by commitment problems?

We extend a recently developed experimental protocol that allows us to measure the impact of shifting power—operationalized by changes in the probability of winning a costly lottery—in a two-stage bargaining game.[7] Our interest here is in

3. Examined from another perspective in Tingley 2017 in this issue.

4. Fearon 1995; Powell 2006. Incomplete information explanations shift the focus to how the anarchic international environment prevents bargainers from revealing their true valuation for a war outcome. Fearon 1995.

5. See Rathbun, Kertzer, and Paradis 2017 in this issue.

6. Hafner-Burton et al. 2017.

7. Tingley 2011.

https://doi.org/10.1017/S0020818316000473 Published online by Cambridge University Press

players' decisions when confronted with conditions of shifting power: do they choose war in the present (as models of commitment problems suggest they should), or accept their opponents' offers and play future rounds under less favorable conditions? Importantly, while war is typically seen as irrational because, *ex post*, it makes both parties worse off, conditions of shifting power change our expectations. In fact, our model demonstrates that when the balance of power is about to shift in your opponent's favor, rejecting offers (choosing war) is strategically optimal.

However, investigating emotions systematically is not a trivial exercise. The difficulty of empirically examining emotions is one of the reasons that we have so little evidence on its effects in IR. Detailed case studies can provide guidance with respect to particular leaders, but are subject to well-known inferential problems because inferring emotions from archival documents can be a tenuous process. Even experiments—often assumed to "fix" many fundamental dilemmas of causal inference—can be limited with respect to emotions, since most studies rely on questions that ask participants how they feel following a "trigger" or emotional stimulus designed to trigger some specific emotion (e.g., anger). This presents problems since individuals (1) often don't have access to their "true" emotions; (2) might not be accurate in recalling how they felt in the past; (3) might not be honest even if they knew how they felt to begin with; and (4) most studies of specific emotions are likely confounded by inadvertently triggering multiple emotions simultaneously but not accounting for their effects in their statistical models.[8]

We address these issues by drawing upon a method novel to IR: the psychophysiological measurement of emotional arousal. This leverages the insight that emotion "rests on a physiological foundation" and "stimulation from the external world results in physiological changes within the human body."[9] In particular, we use skin conductance (electrodermal activity), a physiological measure of arousal in the autonomic nervous system. This allows us to measure participants' sympathetic nervous system activation in a more direct manner than survey-based measures of emotion; one that is free from any potentially disruptive social-masking or impression-management effects, can measure the emotion simultaneous to the stimulus, and bypasses the problems of attempting to stimulate one *and only one* emotion at a time.

The measurement of emotional arousal cannot do everything. It cannot, for example, distinguish between specific emotions like fear or anger. But it allows us to distinguish between the affective and cognitive mechanisms that drive individual decision making in bargaining. This has been a focus in IR for decades, dating back to much earlier debates on the rationality of decision making during crises.[10] A growing body of research in cognitive science has formalized this balance through the framework of the "two-level mode of cognition."[11] Emotionally activated decision

8. Myers and Tingley 2016.
9. McDermott 2004b, 161.
10. Oneal 1988.
11. Hafner-Burton et al. 2017.

https://doi.org/10.1017/S0020818316000473 Published online by Cambridge University Press

processes are strongly associated with intuitive, or "System 1" decision making, and provide a contrast to System 2 which is slower, more effortful decision making characterized by deliberation and conscious thought.[12] While Welch argues that emotion can be a "barrier to good judgment" in foreign policy, we provide evidence on the specific mechanism through which emotions and physiological arousal tamper with our ability to make strategically optimal decisions.[13]

We find that increased emotional arousal leads to a lower probability that individuals take their chances with war in the present, despite that being the optimal solution to their strategic dilemma. We interpret this finding in light of the dual-process model of cognition, which suggests that what we call "aversive arousal" tampers with the ability of those bargaining to act in a strategically optimal manner. This particular finding suggests that emotional arousal is quite a bit more complex than previously suggested: emotions can tamper with the decision process but they don't necessarily have to, and when they do, they may do so in a manner that makes war *less* rather than *more* likely.

Our efforts shed light on the impact of emotions on decision making, but also provide evidence that helps to assess the plausibility of foundational models of bargaining in IR. In accord with the expectations established by our models, we find that larger power shifts lead to higher probabilities of choosing war; individuals prefer to "take their chances" before the power balance turns against them. This fits well with standard explanations of preventive military action in which declining powers prefer to take a gamble *now* rather than face a steep decline relative to their adversaries. Effective institutions can help to mitigate commitment problems: when we allow power to shift but prevent individuals from changing the terms of the bargain across periods, we find that the mechanisms that underlie commitment problems are all but eliminated.

Emotions in Conflict Decision Making

The question of whether and how emotions—and the related dichotomy between intuitive and strategic thinking—affect political decision making is a long-standing puzzle in IR. Because of the obvious and inherent difficulties in "proving" that emotions explain behavior in any given circumstance, what we are left with is mostly speculation. Our goal is to provide micro-foundations that shed light on how emotions might affect decision making in international bargaining, but before we do so, here are several illustrative examples of how emotions and intuition are commonly implicated in international politics.

One class of decisions in which emotions, intuition, and arousal are commonly referenced are those in which leaders react to new and surprising information. For

12. Kahneman 2011.
13. Welch 2003.

https://doi.org/10.1017/S0020818316000473 Published online by Cambridge University Press

example, upon learning that the Soviets had installed missiles in Cuba, President Kennedy famously exclaimed, "He can't do that to me!"[14] Less better known is what his brother, Attorney General Robert Kennedy, said: "Oh shit! Shit! Shit! Those sons of bitches Russians!"[15] More recently, virtually all members of the Bush administration had what were described as emotional reactions upon hearing of the terrorist attacks on 9/11. President Bush's was perhaps the most obvious because he cried on television in the Oval Office and "choked up" in Cabinet meetings in the days following the attack.[16]

Surprise is not a necessary precondition for emotions to be implicated–the stress of a drawn-out conflict can wear leaders down. During the Iranian Hostage crisis, President Carter went through several emotional stages, from "elation" upon learning that Iranian President Banisadr had offered a deal to return the hostages, to bitterness and anger when the deal failed to materialize.[17] In that example, the influence of emotions seems obvious, but in other cases it can lie below the surface. In describing his state of mind in the post-9/11 weeks, CIA Director George Tenet admitted, "I missed my own emotional build-up" and eventually simply "lost it."[18]

Three key issues are apparent in how emotions are implicated in IR. First, there is little to guide us in incorporating emotions more generally into traditional theories of international relations. The conventional wisdom for many years was a version of emotion that was opposed to reason. While this straw man has largely been put to rest, the realization that emotions are neither inherently beneficial nor detrimental for decision making has exposed a large gap in our collective knowledge.[19] We lack well-established theories and empirical foundations for understanding how emotions affect outcomes of interest to IR scholars, leaving us largely in the dark on crucial questions such as whether and how emotions might be incorporated into traditional theories of international relations, as well as under what circumstances emotions are likely to have positive or negative consequences.[20] A number of recent studies have examined the impact of specific emotions, such as trust, attachment, anger, and fear, but such approaches are inherently more narrow and subject to inferential difficulties in distinguishing the effect of similar emotions that are often triggered simultaneously in experiments (such as anger and fear).[21]

Second, in the literature in which emotions are explicitly called upon to help explain outcomes, little evidence is provided for any causal role with respect to international events. Leaders may indeed have appeared to have emotional breakdowns, but even if we grant that basic fact (which requires a leap of faith given the difficulty

14. Quoted in Allison 1969, 713.
15. Quoted in Stern 2005, 22.
16. Woodward 2002.
17. Harris 2004.
18. Hall and Ross 2015, 864.
19. On emotions as a component of rationality, see McDermott 2004a.
20. For an overview, see Renshon and Lerner 2012.
21. On the difficulty of disentangling specific emotions, see Myers and Tingley 2016.

https://doi.org/10.1017/S0020818316000473 Published online by Cambridge University Press

in imputing emotions to others, incentives to misrepresent, etc.), it is far more difficult to argue that this *caused* any particular event to occur. While statesmen typically overestimate the influence of individuals, political scientists are justifiably skeptical that individual reactions are anything more than noise distracting us from the larger, structural forces at work.

A final issue relates to the overlapping Venn diagram one might draw to describe the literature on emotions in IR. While a proliferation of recent work has made important advances, important gaps remain. For example, one can easily find work bearing on either individual or collective emotions, a focus on individual leaders or the emotions of normal citizens, and on biological and physiological mechanisms and more abstract notions of "affect." However, among the subset that focus on decision making, approaches and standards of evidence vary tremendously, from traditional case studies to those that argue that "research can be insightful and valid even if it engages unobservable phenomena, and even if the results of such inquiries can neither be measured nor validated empirically."[22]

Without disparaging previous efforts, we note that there is little to guide one in searching for systematic, causally identified evidence on the impact of emotions—and the physiological mechanisms through which they operate—on decision making in IR. To begin this process, we use bargaining theory as a foundation on which to build our empirical inquiry into the effects of emotions in IR. Doing so provides the significant advantage of allowing us to begin from a body of well-developed and precise theoretical predictions. This in turn means that we are able to cleanly identify whether and how emotions affect decision making by causing individuals to deviate from the predictions generated by our formal model of bargaining.

Bargaining, Shifting Power, and Physiological Arousal

Commitment problems have gained traction as one of the more powerful explanations of conflict in the study of international relations. They are part of a body of literature that identify "mechanisms" that lead to conflict. In the sense the term is most often employed, *mechanisms* denote a set of utility-maximization assumptions and a particular bargaining environment that result in predictions about when conflict will occur. However, these models do not offer guidance on how *individual decision makers* might be affected by the bargaining environment, nor how they are likely to respond. This is a critical next step in understanding the dynamics of international conflict, just as the study of micro-level behaviors has furthered our understanding of strategic choice throughout social science and in IR specifically.[23]

22. Bleiker and Hutchison 2008, 115.
23. Ostrom 1998; Tingley and Walter 2011b.

https://doi.org/10.1017/S0020818316000473 Published online by Cambridge University Press

A Model of Commitment Problems

The intuition behind this type of bargaining model of conflict is as follows.[24] There are two decision makers from competing countries (Argentina and Brazil, or more generally, A and B) and two periods (present and future). In Period 1 (present), Argentina makes an offer to Brazil concerning the division of a resource that has a value of 1. If that offer is accepted, then the game continues to Period 2 (the future) and Argentina again makes an offer. As before, if Brazil accepts the demand, the resource is divided in the specified manner and the game ends.

If in Period 1, Brazil *rejects* Argentina's offer, then the two players play a war lottery in which one player wins the resource (with probability p) but both players pay some costs (since war destroys some of the value of the stake). If a "war" is fought in Period 1, then the winner automatically gets the whole resource in Period 2 (the future). If Argentina's offer is rejected in Period 2, then (as before) both states play a war lottery to determine who wins the resource (less the costs of the war).

The key question in this model is: what will make Brazil reject the offer and choose war in the present (Period 1)? If Argentina's power is increasing relative to Brazil's, then a commitment problem dynamic exists, since there is no way for Argentina to credibly commit to not taking advantage of Brazil in Period 2 (by demanding most or all of the resource). We can operationalize this shifting power dynamic by increasing the probability of Argentina's victory in Period 2. If this dynamic holds, one would expect Brazil to make the strategically optimal decision to choose war (i.e., reject any offer made by Argentina) in Period 1, since doing so allows it to take its chances with a war lottery before the power dynamic has shifted in favor of its adversary. This is the logic of preventive war.

If, on the other hand, both states expect the power balance to remain stable across periods, we wouldn't expect the dilemma posed by commitment problems to be as severe, and we would thus expect to see fewer instances of preventive war. In the studies we describe we "activate" the commitment problem dynamic by making A's probability of winning increase between Periods 1 and 2 of the game. We also present a follow-up experiment inspired by Quek that allows for a large shift in power, but no ability to change offers in Period 2. That is, A actually can *credibly* commit to not taking advantage of its adversary in the future since both A and B know that A's offer cannot change between periods.[25]

Notice that this model does not offer mechanisms particular to the human decision makers that are ultimately of interest, nor does it incorporate factors (such as emotions) that seem likely to affect the judgment of the actors involved. To see this clearly, consider the contrast between the pioneering work in behavioral economics that seeks to understand the physiological bases of strategic choice versus scholarship

24. We provide a more detailed exposition of this model in online Appendix A.
25. Quek 2017.

https://doi.org/10.1017/S0020818316000473 Published online by Cambridge University Press

that relies on the abstract choice models themselves.[26] What physiological mechanisms operate when people face situations with commitment problems? Do individuals calmly calculate the prospects for conflict and respond accordingly, or do automatic, emotional processes become involved?

Dual Process Theory: A Window into Emotions in IR

The idea that individuals' cognitive processing can be described as two parallel systems can be traced back at least as far as William James and Sigmund Freud, who described a dual system of "rational" and "irrational' thinking.[27] We follow Evans and Over in conceptualizing two parallel cognitive processes, System 1 and System 2.[28] The former is based upon previous experiences and beliefs and is characterized as "implicit, associative, fast and highly robust."[29] Perhaps most importantly, it is automatic: it occurs all the time without our awareness. In contrast, System 2 is slow, controllable, and explicit (in the sense that one is consciously engaging in the act of analytical reasoning). System 1 is, in Haidt's analogy, the elephant, while System 2 is the rider; the rider can provide advantages over pure intuition by examining alternative scenarios, thinking into the future, reasoning backwards, and acting as spokesman for the elephant (without truly knowing what the elephant is thinking).[30] However, the rider is not always in control and while our automatic processing has been honed by millions of years of evolution, it is also particularly vulnerable to the "catalog of biases" now documented in behavioral science.[31]

This model of cognition provides a window into affective dynamics since emotions are more closely associated with System 1 processing. That is, individuals seem to feel an emotion, and then construct a rational edifice around that feeling to explain why they are angry, sad, happy, etc.[32] Emotions have been explicitly linked to System 1 processing with evidence for this fact often based upon the parts of the brain involved in both emotion processing and System 1 processes.[33] That emotions are so strongly associated with System 1 processing is of immense help empirically because we are on firmer ground in measuring the generalized physiological arousal associated with System 1 than we are in targeting specific emotions using experimental designs.[34]

26. For a brief review, see Camerer 2011, chapter 2.
27. See Osman 2004. For an overview of some of the controversy associated with the framework, see Gigerenzer and Regier 1996.
28. Evans and Over 1996. See also Kahneman 2011.
29. Gilinsky and Judd 1994.
30. Haidt 2012.
31. Hafner-Burton et al. 2017.
32. Haidt 2001.
33. See Evans 2008.
34. Renshon, Lee, and Tingley 2015; Myers and Tingley 2016.

https://doi.org/10.1017/S0020818316000473 Published online by Cambridge University Press

How do these different processing styles—with their differential reliance on emotions—affect bargaining? As we alluded to earlier, most modern understandings of emotions have moved past a simplistic dichotomy in which emotions are considered irrational and bad in contrast to "cool cognition" and "logical" processing of information. In fact, in nonpolitical contexts, there is a substantial amount of research on the benefits of System 1 processing.[35] In IR contexts, however, it is far more rare for scholars to make the argument that the affect-laden System 1 processing might be beneficial. One notable exception is Blight, who argued that the fear US leaders experienced at the height of the Cuban Missile Crisis was adaptive in aiding their careful navigation of the crisis.[36] McDermott offers a more theoretical treatment by making the important point that emotional processing is not definitionally opposed to rationality, and in fact can be a key component of making strategically optimal decisions.[37] Without emotions, we are "stuck, more fatally than Buridan's ass, who died of exposure to two equidistant carrots."[38]

However, despite decades of research in related fields demonstrating that emotion is critical for rationality, an underlying assumption of much of IR remains that "logical, rational calculation forms the basis of sound decisions."[39] Because there is so little empirical research on this question, we explain how processing styles may affect political decision making and then draw hypotheses from related behavioral work on bargaining.

The first and most obvious way that processing style might have an effect on decision making is that political leaders might act on their initial emotional reaction to a situation. In high-pressure, time-sensitive crises, there may simply not be very much time for reflection or deliberation. Emblematic of this is the "explosive reaction" of Kissinger and Nixon to Syrian intervention in Jordan in 1970.[40] And in general, the time pressure associated with intense crises seems to exacerbate the influence of emotions.[41]

However, even if there is abundant time to reassess the situation and deliberate over options, the initial automatic response may be quite important. What Holsti called the "definition of the situation" and what psychologists refer to as "framing" may exert tremendous influence over how the situation is handled even after reflection.[42] The reason for this is twofold. Even in longer time frames, it is often the case that some options are discarded immediately so that initial reactions may have far-reaching effects by limiting the choice set of actions, even if the remaining options are considered carefully. Further, individuals often use analogies and metaphors

35. Gigerenzer 2007.
36. Blight 1992.
37. McDermott 2004a.
38. De Sousa 1987, 15. Technically, we believe the donkey was between a stack of hay and a pail of water, not two carrots.
39. McDermott 2004a, 691. On emotions and rationality, see Damasio 2008.
40. Dowty 1984, 153.
41. Crawford 2000.
42. Holsti 1970.

https://doi.org/10.1017/S0020818316000473 Published online by Cambridge University Press

that are chosen instinctively to frame their understanding of an event.[43] If we return to the well-trod territory of the Cuban Missile Crisis, one notices that even though Kennedy is often congratulated by scholars for engaging in a thorough, deliberative decision-making process (following his initial, emotional reaction), the basic premise of the crisis—particularly that it *was* a crisis—was settled on instinctively and never once challenged.

Hypotheses

In our experimental set-up there are two periods and two actors (a proposer and a responder). In Period 1, the proposer offers some division of resources; if the responder accepts, the resource is divided as proposed. However, the responder's rejection leads to a gamble where one of the players receives all of the resources. In the declining power (or *BigShift*) condition, the responder wins all of the resources with a probability of .7 in Period 1 and .3 in Period 2 (the probabilities are reversed for the proposer, who wins with a probability of .3 in Period 1 and .7 in Period 2). In the control (or *SmallShift*) condition, the responder's probability of winning declines, but only very slightly, staying at roughly .5 for both periods.[44]

In other words, in the declining-power (*BigShift*) condition, both parties know that the proposer will be gaining power in Period 2 (the future), and the responder will be losing power. Given this knowledge, the rational course of action for the responder is to reject *any* initial offer and take their chances in a costly lottery in Period 1, when they still enjoy a power advantage (operationalized as a high probability of winning the costly lottery for the resource). Thus, responders in the *BigShift* condition should, if motivated by self-interest and using rational tools of backwards induction, reject all offers, even high ones. This, at last, is what standard models of IR would predict in the absence of any other factors. We operationalize this by investigating whether—for any given offer size—there will be a higher probability it is rejected when there is a large power shift.

While H1 describes our expectations in the absence of any other factors, a wide body of research suggests that other factors, including emotional arousal, *are* present. Initial evidence from studies on "ultimatum games" (UG) suggests that emotions play a powerful role, despite assumptions that strategic thinking or "cold calculation" should dominate.[45] For example, one study found that low offers (made by a

43. Khong 1992.

44. Responder wins with probability .55 in Period 1 and .45 in Period 2.

45. Ultimatum Games are similar to our present study only in that there are two players—a proposer and a responder—and the task is to divide resources. In the UG, one person proposes a division of a resource to a second person. If the second person accepts, then the resource is divided in the agreed-upon manner; if the responder rejects the offer then neither person gets anything. Thus, the responder is always better off accepting as long as the offer >0. In other words, the standard equilibrium analysis of the UG is that there should never be rejected offers, even when they are arbitrarily low. Our game is considerably

https://doi.org/10.1017/S0020818316000473 Published online by Cambridge University Press

proposer) resulted in responders exhibiting a higher level of activation in their bilateral anterior insula, an area of the brain associated with negative emotional arousal.[46]

However, "low" offers should carry different implications depending on the situation's context and structure. In fact, the link between bargaining and emotional arousal seems to be dependent more specifically on offers that are considered unfair or threatening in some way. In one UG study, low offers provoked anger in responders, but only when the allocation was intentional and not when both parties knew it was randomly assigned.[47] Van't Wout and colleagues and Civai and colleagues similarly observed increased electrodermal activity (skin conductance), which is associated with insula activity, when responders received unfair offers.[48] Thus, emotions like anger can be triggered under certain circumstances, but there is a logic to this process; a "bad" or "low" offer in a bargaining context is, by itself, not enough to trigger an emotional reaction. There must be some knowledge of the motivation behind the allocation (to know whether there was an *intent* to be provocative or threatening) or a larger understanding of the amount available to the proposer (to know if the allocation was fair).

We operationalize emotional arousal as physiological arousal, measured as skin conductance levels (SCL). SCL is a reliable indicator of increased attention and arousal and is associated with greater insula activities.[49] Skin conductance has also been used in similar contexts to investigate the relationship between individuals' emotions and their decisions.[50] Physiological arousal does not allow us to distinguish the valence of our subjects' feelings (i.e., positive or negative feelings) or disentangle specific emotions from one another. However, our theoretical focus is on the intensity of emotional arousal, and contains no specific predictions for what the effects of valence (or specific emotions) might be. Given that one's autonomic nervous system responses precede conscious awareness and are difficult to consciously control, we believe that the fluctuation (i.e., reactivity) in skin conductance should provide valuable information about the extent to which intuitive versus deliberate processes are being utilized.[51]

Following previous work, we thus expect that low offers will lead to higher levels of physiological and emotional arousal, *but only in the presence of a large power shift*. In that context, when subjects expect to decline in power in the future, low offers might be more likely to be taken as a signal of threatening intent. This follows from the logic that the presence of a shifting power dynamic is likely to make lower offers more psychologically aversive than in conditions where power

more complicated because it adds conditions of shifting power, a costly lottery, and multiple periods. Equilibrium predictions thus differ significantly between the two experimental set-ups.

46. Sanfey et al. 2003.

47. Blount 1995.

48. Van't Wout et al. 2006; Civai et al. 2010.

49. See Sequeira and Roy 1997.

50. Van't Wout et al. 2006; Civai et al. 2010. In political science contexts, see Mutz 2007; Renshon, Lee, and Tingley 2015.

51. Gardner, Gabriel, and Diekman 2000.

https://doi.org/10.1017/S0020818316000473 Published online by Cambridge University Press

remains stable across periods. In other words, the proposer's lower offer in Period 1 will be interpreted as more threatening and challenging to the responder, knowing that the proposer will be able to take advantage of increased bargaining power in the next round. On the other hand, small shifts should not activate the same aversive physiological responses because the responder is not as likely to be concerned with future bargaining.

As for the *consequences* of that emotional arousal, current research on emotion provides mixed guidance. Even if we move beyond a simple dichotomy between emotion and cognition, it is not obvious how arousal might affect decision making in our bargaining context. Just because emotions are not inherently opposed to rationality does not mean that they uniformly improve decision making in every context and at every level of intensity. In fact, Mcdermott noted that "even if emotional information *usually* adds accuracy and efficiency to judgment and decision making, *extreme emotions may limit the ability of decision makers to assess a situation more objectively*."[52]

The results of several economic bargaining studies suggest that while emotions may be helpful in some cases, they do not seem to improve our ability to think strategically. In these situations, emotional (and thus physiological) arousal seems to tamper with a more deliberate, cognitive processing of information. According to the dual-process framework, decisions are made by the interaction of intuitive and reflective processes.[53] Emotional and physiological arousal give rise to more intuitive decision making, which is associated with fast, automatic, and effortless processing.[54] In a landmark review of emotions, Loewenstein and Lerner noted that "at sufficient levels of intensity, emotions can overwhelm cognitive processing and deliberative decision making altogether."[55]

Conversely, the absence of physiological arousal allows for more reflective decision making, which is associated with slow, effortful, and deliberate processing. For example, delays (in effect, a "cool down" period) in one bargaining study allowed respondents time for more deliberation and in turn increased the probability that players accepted offers.[56] In other studies we noted earlier, higher physiological/emotional arousal (as measured by activation of the insula region of the brain) was significantly correlated with non-equilibrium responses.[57]

In our bargaining game, the presence of large power shifts implies that the responder's dominant strategy is to reject any offers in Period 1: the responder's chances in Period 1 will *always* be better than in Period 2 under conditions of declining power. In thinking through this dilemma, it is easy to imagine oneself in the responder's position. If the proposer makes a low offer, I should reject that offer, either because I'm

52. McDermott 2004a, 700–701, emphasis added.
53. Kahneman 2003; Sloman 1996.
54. Lambourne and Tomporowski 2010.
55. Loewenstein and Lerner 2003, 627.
56. Even when accepting them was the optimal response; see Grimm and Mengel 2011.
57. Sanfey et al. 2003.

https://doi.org/10.1017/S0020818316000473 Published online by Cambridge University Press

sophisticated (and know that my odds will be better now than they will be later), or because the low offer makes me angry (which will spur me to attack).

Our prediction runs opposite to these two intuitions, and it is worth explaining why. Of course, as social scientists familiar with bargaining theory, we know that players' odds are better in Period 1 than they will be in the future. Actually choosing this strategy, however requires the ability to look ahead and induct backwards and as Camerer argues, people's ability to engage in such strategic, iterated thinking is limited.[58] Our argument is that emotional arousal is likely to short-circuit players' ability to think strategically and make the "correct" choice.

The second and related issue is why the low offer (in the *BigShift* condition) wouldn't simply make the player angry enough to lash out and choose the war lottery. In our experimental set-up, doing so is actually the strategically optimal path, so if anger were to lead to this choice (because subjects, in their spite, want to destroy shared resources), we might describe this as doing the "right thing for the wrong reasons." However, we see this as relatively unlikely since whatever physiological/emotional arousal we measure is not necessarily anger, and might well be one or some combination of anger, fear, or anxiety. All of these different emotions carry different behavioral implications; while anger might make subjects lash out, fear might lead to the opposite reaction. We are thus on much safer ground specifying the way in which intense emotional arousal—regardless of which *particular* emotion is evoked—might distort decision making.

We thus predict that in the *BigShift* condition, lower physiological arousal will be associated with strategically optimal play. In this context, that means that responders will reject offers in Period 1 and take their chances before their relative power declines. On the other hand, we predict that high physiological arousal should interfere with this process of strategic thinking, which requires sufficient cognitive resources to think about the implications of the shifting power dynamics that they will face in Period 2. Responders who experience intense physiological arousal will thus be more likely to take their chances in Period 2, despite facing worse odds in the future.

In the formal causal mediation framework that we use to analyze our data, this can be stated as follows. We predict that the effect of low offers on rejection rates that is transmitted though higher physiological arousal will be positive; that is, it will lower rejection rates and by implication increase the probability of *non*-optimal choices. Conversely, larger offers, even in the presence of a power shift, will not trigger physiological arousal, allowing players to analytically dissect their situation and choose the optimal strategy (rejection of *any* offer). Higher (better) offers should thus decrease physiological arousal, which in turn will lead to higher rates of rejection (the strategically optimal choice).

Our final hypothesis concerns the impact of institutions on ameliorating the commitment problem. The dynamics of the commitment problem as we have described them suggest that low offers from actors rising in power are psychologically aversive;

58. Camerer 2003.

https://doi.org/10.1017/S0020818316000473 Published online by Cambridge University Press

they appear threatening to actors declining in power who are likely to see it as a signal that they will be preyed upon in the future. However, if the effects of the power shift are mitigated, we hypothesize that low offers from actors rising in power are apt to seem much less threatening. Thus, in scenarios where institutions were able to guarantee that rising powers would not prey on the weaker power in the future, aversive (low) offers from the rising power should not generate the same psychological or physiological reaction that such offers would have without those institutions.

H1: Larger power shifts will lead to higher probability of rejection for any given level of offer made by proposer.

H2: Lower offers should be associated with increased physiological arousal (higher skin conductance levels), but only in the BigShift *condition.*

H3: When there is a large power shift, the effect of offers on rejection rates that is mediated by heightened physiological arousal will be positive. This mediation effect will be absent in the low power shift condition.

H4: When offers are not allowed to change in between periods, heightened physiological arousal will no longer mediate the relationship between offer size and the decision to accept or reject.

Experimental Design

Lab Procedure

We conducted this study at the Harvard Decision Science Laboratory using 121 male undergraduate students ($age_{mean} = 22$) who signed up. Each session began with subjects being hooked up to the physiological sensors we will describe. Then subjects read a set of instructions that explained the basic rules of the bargaining game using neutral language. A condition- and hypothesis-blind experimenter also gave the instructions verbally and then answered any questions. After obtaining a baseline recording of the skin conductance (since physiological *reactivity* is measured against a baseline, resting rate for each individual subject), the bargaining game began.[59]

Each subject was randomly paired with another subject and randomly assigned a position (proposer or responder). Subjects then played the bargaining game sequentially. At each stage of the game a unique signal was sent to the physiological recording device to indicate different outcomes in the game, such as an offer being received,

59. Subjects watched a video (just under three minutes long) featuring relaxing images of beaches and palm trees with calm music in the background to measure the physiological responses at the baseline. This particular stimulus and protocol has been used in previous research, and by the authors in Renshon, Lee, and Tingley 2015.

https://doi.org/10.1017/S0020818316000473 Published online by Cambridge University Press

through parallel port communication. This enabled us to precisely match a given subject's physiological response at a given time to specific actions in the bargaining game. After playing ten matches (of two periods each), subjects took a short postexperiment questionnaire that included several policy questions and personality inventories. Subjects were paid privately based on one randomly selected match. On average, subjects earned $20 for approximately fifty minutes of time.

Subject Population

There are trade-offs of our chosen method.[60] Regarding the use of undergraduate subjects in a laboratory experiment, we note two critical points. First, while elite experiments are becoming somewhat more common, they are logistically implausible in many circumstances.[61] Obtaining elite samples is costly in time and resources and often not possible without special connections that few researchers have. Additionally, when possible, most scholars design elite studies to maximize participation rates, which often means relying on survey instruments.[62] Thus, even among the subset of experimental studies that use elites, laboratory studies are rare.[63] This is a critical point, since some types of studies are simply not possible outside of a laboratory, including the present one with its focus on measuring emotions using psychophysiological tools.

A broader and more important point (than logistical concerns) is that the primary rationale for elite experiments *must be* that leaders/elites are hypothesized to differ systematically from convenience samples on dimensions that are theoretically relevant to the study at hand. In our context, one would have to argue that, for example, elites would be less likely to have emotional responses. Neither anecdotal nor systematic evidence that we can find suggests that this would be the case. If anything, emotionality (and the behavioral biases with which it is associated) *increases* with political sophistication.[64] Thus, we contend that—without strong prior beliefs that a theoretically relevant difference exists between elites and students and that we know the direction of the resulting bias—the benefits of the laboratory for studying emotions are worth the potential trade-off.

60. For a longer discussion of these issues, see Renshon 2015.

61. For examples of elite experiments, see Hafner-Burton, Hughes, and Victor 2013; Renshon 2015; Renshon, Yarhi-Milo, and Kertzer 2015. In fact, the proportion of elite versus convenience samples in this issue is likely representative of the field more broadly. Kertzer 2017; Rathbun, Kertzer, and Paradis 2017; and this study use undergraduate subjects in a laboratory experiment. Herrmann 2017 uses nationally representative samples and a survey instrument; Tingley 2017 uses a combination of nationally representative and MTURK samples and a survey instrument; Rho and Tomz 2017 use an MTURK sample and a survey instrument; and one study (Bayram 2017) uses a survey instrument fielded on a (small) sample of German politicians.

62. For example, Loewen et al. 2014; Renshon, Yarhi-Milo, and Kertzer 2015; Bayram 2017.

63. Though they do exist, see Hafner-Burton, Hughes, and Victor 2013; Renshon 2015.

64. Miller 2011.

https://doi.org/10.1017/S0020818316000473 Published online by Cambridge University Press

A more general concern is that even if we had access to subjects of theoretical interest, the artificiality of the lab prevents us from generalizing to situations in the White House Situation Room. While lab experiments are in some sense purposely artificial, the value gained from the trade-off is significant: an ability to isolate concepts that otherwise are nearly always inferred in a tautological manner from observational data.[65] To the extent that the artificial nature of experimental designs is problematic, we are nearly always better served by thinking of such experiments as part of larger bodies of evidence compiled from different data sources and different research designs.[66] We also note a long history of laboratory studies being replicated in diverse environments, using different samples, measures, and contexts.[67]

The common criticism of IR experiments that the stakes in the real world are so high that experiments cannot mimic them rests on shaky foundations. In fact, empirical work on this subject is relatively clear: stakes do not change the dynamic dramatically in most instances, and to the extent they do, it is the inclusion of moderate stakes (compared to no financial incentives) that make a difference, not marginally increasing stakes that are already present in the experiment.[68] This is why our study was designed so that subjects' decisions do have real (if small) economic consequences; participants lose *actual* money if they perform poorly, so there are strong incentives to pay attention and try to "win."[69]

Experimental Game and Treatments

In our experiment, subjects were anonymously paired with each other and randomly assigned to be in position A or B. After two periods of the game, subjects were randomly matched with another subject and played another two-period game (and so on, for ten total matches). Each period of each match had the same basic structure. In Period 1, player A proposes a division of a resource of size $R = 10$, $(x_{1A}, x_{1B} = 10 - x_{1A})$. The subscripts represent the period and position, so x_{1B} is what player B would get from player A's first-period offer. Player B decides whether or not to accept the offer. If they reject (i.e., choose a war lottery), the entire resource is assigned by the computer to player A with probability p_1 and B with probability $1 - p_1$, where the subscript again denotes Period 1. Both players also pay a cost c_A, $c_B = 2$, destroying some of the total value of what is being fought over. If B

65. See discussion in Kertzer 2017.
66. As in Rathbun, Kertzer, and Paradis 2017.
67. See Levitt and List 2007.
68. See Holt and Laury 2001. In fact, the relationship between incentives and performance does not appear to be entirely monotonic when incentives are the extremely high, suggesting that ratcheting up the stakes is not necessarily an appropriate fix. See Ariely et al. 2009.
69. On the other hand, Kahneman and Tversky 1979 recommend large hypothetical gambles over "contrived gambles for small stakes." Notice that this solution would be even more unacceptable to those political scientists worried about external validity than the inclusion of "only" small stakes. This may say more about our almost religious beliefs concerning the appropriateness of various methods than anything else.

https://doi.org/10.1017/S0020818316000473 Published online by Cambridge University Press

chooses to reject the offer in Period 1, the war lottery occurs and the winner receives $R = 10$ points automatically in Period 2. If B accepts the offer in Period 1, the game moves on to Period 2, where another $R = 10$ points are divided and both players stay in the same position (A or B).

To activate the commitment problem, we used two experimental conditions that mimicked a shifting of the balance of power by changing the probability that each player would win the resource should a war lottery occur. In *BigShift* $p_1 = .3$ and $p_2 = .7$. In other words, player A starts out "weaker" in Period 1, but gains a significant amount of power (i.e., higher probability of winning the costly lottery) so that in Period 2 the power dynamics have reversed. In this situation, our game-theoretic model predicts that player B should reject *any* offer in Period 1, since "taking their chances" in a Period 1 war lottery is preferable to doing so in Period 2 once a major power shift has occurred. This situation mirrors the commitment problem dynamics described earlier: player A has no way to credibly commit to not taking advantage of their increased power in the future and player B thus has strong incentives to fight for the resource while their position is advantageous.

In our second experimental condition, *SmallShift*, $p_1 = .49$ and $p_2 = .51$. In both conditions, the power shifts in the same direction (player A becomes more powerful in Period 2), but in *SmallShift* the magnitude of the change is negligible and thus unlikely to be associated with commitment problem dynamics. In this condition, our model predicts fewer rejected offers and a higher likelihood of a "peaceful" bargaining resolution as a result of the lack of a commitment problem. In particular, the model predicts that any offer over 5.3 should always be accepted, whereas in the *BigShift* condition all offers should be rejected (see online appendix for additional details).

All aspects of the game are common knowledge. Subjects were paid based on one randomly selected repetition of the experiment and received one dollar for every ten points earned.

Physiological Data Acquisition

Participants' skin conductance levels (SCLs) were measured continuously throughout the experiment. We followed standard procedures for acquiring and scoring the skin conductance data (details can be found in online Appendix B). To address potential individual differences in variability in skin conductance, our data were transformed to deviations from each participant's baseline and standardized within each participant. Because our experimental design required a measure of physiological *reactivity*, we needed something meaningful to compare to our baseline levels of skin conductance for each participant. We did this by comparing the physiological arousal of participants *while they made their decision*; that is, *after* a proposer made an offer and *before* responders made the decision to accept or reject the offer in Period 1. The average deviation from baseline during this particular phase of the experiment is our key physiological variable. Practically, this was made feasible through our development of computer protocols that placed "tags" in the physiological data at signals

https://doi.org/10.1017/S0020818316000473 Published online by Cambridge University Press

(transmitted via parallel port communication) from the bargaining game operating on another computer.

Results

Figure 1 presents histograms of the amount offered to Player B in Period 1, with vertical lines indicating the average offers in the experimental conditions. Average offers, and variations in offers, were both higher in the *BigShift* condition were slightly higher.

Do Power Shifts Create Commitment Problems?

A key implication from our formal model was that power shifts should "activate" commitment problems in our bargaining scenario. To that end, we expected to see that Player B was more likely to choose the war lottery (i.e., reject A's offer) when a power shift was imminent (*BigShift*) than when it was not (*SmallShift*).

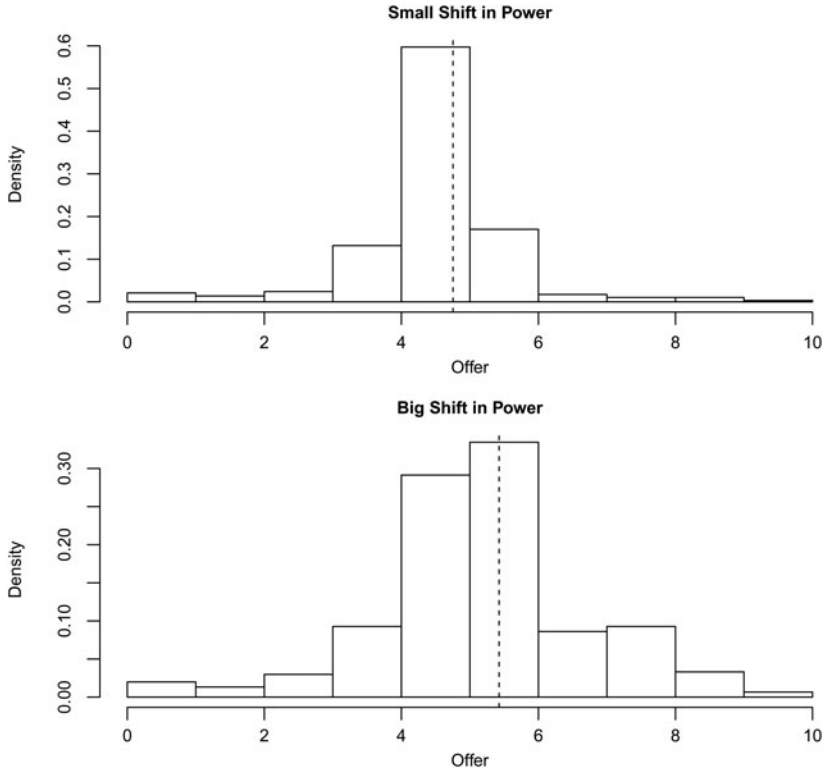


FIGURE 1. *Histogram of first-round offers by power shift condition*

https://doi.org/10.1017/S0020818316000473 Published online by Cambridge University Press

We estimated a probit regression model in which the dependent variable was whether or not the offer was rejected. The independent variables were the size of the offer, an indicator for the experimental condition (*BigShift* or *SmallShift*), and an interaction between the experimental condition and offer size. Standard errors clustered at the individual level. Using the results of this model, we then plot the predicted probability of an offer being rejected in Figure 2 using the Zelig package in R.

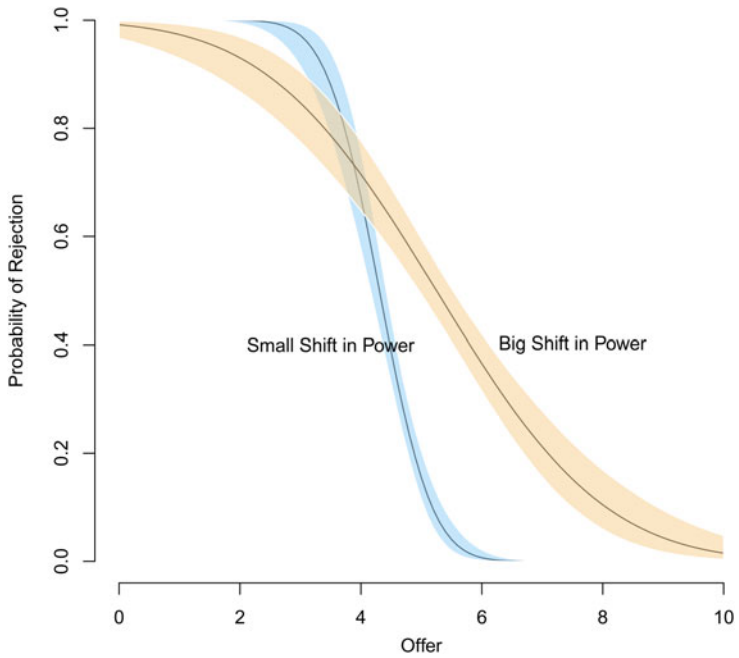


FIGURE 2. *Probability a Period 1 offer is rejected as a function of offer size and experimental condition*

Our results were consistent with H1: offers are more likely to be rejected in the *BigShift* condition. Offers below 5 were almost always rejected in both the *BigShift* and *SmallShift* conditions. However, above an offer size of about 5, offers are more likely to be rejected when in the presence of a large power shift. Thus, across both conditions, individuals are more likely to accept generous offers than stingy offers. However, the presence of a power shift does affect the rate at which generous offers were rejected.

Do Offers Trigger Physiological Arousal?

Next we investigate the impact of the size of Player A's offer on Player B's emotional arousal. Our prediction was that offer size would be negatively correlated with

https://doi.org/10.1017/S0020818316000473 Published online by Cambridge University Press

physiological reactivity (better offers would lead to lower arousal), and that this would be more apparent in the *BigShift* condition. We estimated a simple linear regression model with our measure of physiological arousal as the DV and offer size as the IV.[70] We found that higher offers on average lead to lower levels of skin conductance, but only in the *BigShift* condition ($t = -2.19$, $p < .05$). In the *SmallShift* condition there was no significant relationship between offer size and arousal ($t = -.3$, *n.s*). To display the results, Figure 3 plots predicted arousal as a function of offer size.

This finding suggests that responders in the *BigShift* condition were more physiologically reactive to low offers than to high offers, as we predicted. To explain this, we must recall the context in which Player B is operating. In the *SmallShift* condition, neither high nor low offers should trigger physiological arousal, since B can simply take their chances with a war lottery in which they have roughly even odds of winning. In the *BigShift* condition, large offers in Period 1 must seem attractive, and B might even interpret these as an attempt by A to compensate them for the

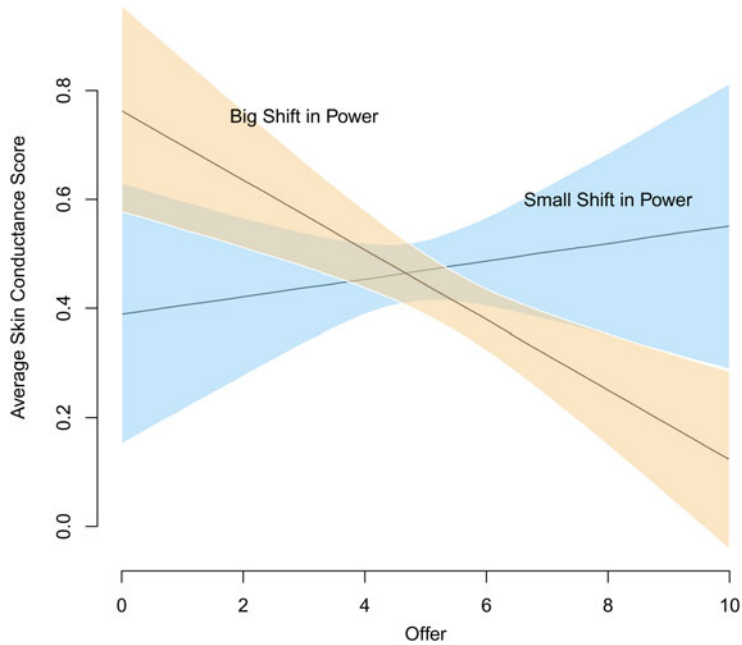


FIGURE 3. *Effect of offer sizes on skin conductance. 90% confidence intervals used for presentational purposes and robust standard errors*

70. We also estimated generalized additive models that do not impose a linearity assumption. Results were nearly identical to the linear model.

https://doi.org/10.1017/S0020818316000473 Published online by Cambridge University Press

risk they face in Period 2 when B will have less leverage. Low offers in the *BigShift* condition, however, are perceived as provocative, and may thus be interpreted as a challenge or signal of ill intent, which might have activated more physiological arousal among the responders.

Do Offers Affect Decisions Through *Physiological Arousal?*

Finally, we turn to the role of physiological arousal in mediating the relationship between offer size and decision to choose the war lottery or take one's chances in Period 1.[71] We estimated a mediation model based on two parametric models. The first is the model used earlier which regresses physiological arousal on offer size. The second model is a probit regression in which a binary variable indicates whether the offer is rejected (1) or accepted (0). We also permit an interaction between the treatment and mediator in the outcome model, allowing for $\delta_i\ (t) \neq \delta_i\ (t')$. This could occur if the postulated arousal process occurs only when offers are larger, rather than smaller. We estimate these models separately for the two experimental conditions (*BigShift* and *SmallShift*).

Using the results from these models, we estimate the mediation effects using the methods described by Imai, Keele, and Tingley.[72] With this method we must first define the treatment contrast. In our setting, this amounts to setting two treatment conditions, t', t, that correspond to two different offer sizes. We focus on the contrast between $t' = 5$ and $t = 10$. Thus, for example, in the second contrast $\delta_i\ (10)$ represents the change in probability of rejection that occurs with an offer of the entire resource that is a result of the difference in physiological arousal ($Y_i\ (10, M_i\ (10)) - Y_i\ (10, M_i\ (5))$). A positive change in probability is consistent with our H3. When there is a large power shift, the effect of offer size on rejection rates that is mediated by heightened physiological arousal will be positive. Put differently, this prediction states that low offers will (through the heightening of physiological arousal) decrease rejection rates and by implication *increase* the probability of *non*-optimal choices in the *BigShift* condition. This would be consistent with our theory of "aversive arousal" in which intense physiological/emotional arousal short-circuits the process of deliberation and reflection. In contrast, because low offers are unlikely to be perceived as threatening in the absence of a power shift, this mediation affect should be absent in the *SmallShift* condition.

Figure 4 plots the result for the *BigShift* (top) and *SmallShift* (bottom) conditions. In the plot, the average causal mediation effect (ACME) for both treatment conditions (dark line for t and dashed line for t'), direct effect, total effect point estimates, and confidence intervals are reported. The mediation effect under t (dark, solid line) is

71. A discussion of the assumptions required of mediation analysis can be found in online Appendix C.
72. Imai, Keele, and Tingley 2010. Because a probit model is used, common approaches such as the product of coefficient method/Sobel test are not applicable. We obtain substantively identical results if we use a linear model for the rejection choice.

https://doi.org/10.1017/S0020818316000473 Published online by Cambridge University Press

positive and statistically significant for the *BigShift* condition while the mediation effect for an offer of 0 is not different for the *BigShift* condition, nor is either mediation effect statistically different from 0 for the *SmallShift* condition. Consistent with H1 and the previous results, the total effect of offer size on rejection decisions is negative, large, and statistically significant (in both *BigShift* and *SmallShift* conditions). Overall, lower offers led to a higher probability of rejection: individuals in our experiment seem to prefer more money than less, an effect present in both conditions.

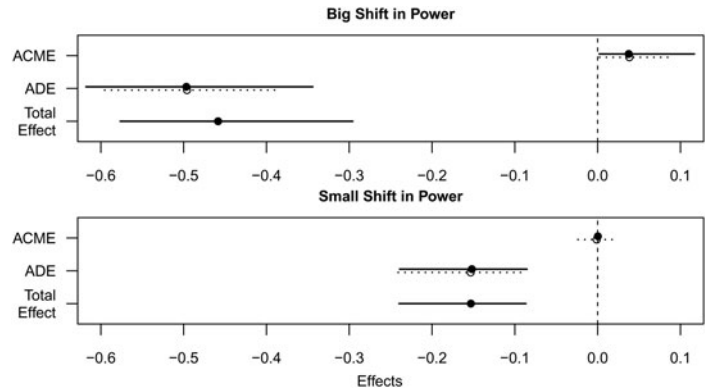


FIGURE 4. *Mediation Effects for treatment change from 5 to 10. Top row is* BigShift *condition and bottom row is* SmallShift *condition, with 95% confidence intervals allowing for clustering at the individual level*

Higher levels of physiological arousal thus lead to a *lower* probability that the player chooses "war" (i.e., rejects offer). Put differently, physiologically aroused subjects chose to accept low offers at a greater rate than their "less aroused" counterparts (a finding robust to controlling for offer size). On the face of it, this may perhaps seem puzzling; our intuitions typically lead us to associate "emotional arousal" with aggression and conflict, but here the relationship is reversed. Here, physiologically aroused subjects were *less* likely to choose war, and more likely to accept Player A's Period 1 offer. How do we reconcile this seeming contradiction? The key is in the realization that if Player B chooses war, they choose the *strategically optimal* path. That higher levels of arousal are associated with the opposite—accepting their adversary's offer and taking their chances in a second round in which their own power will be greatly diminished—suggests that the arousal is preventing sound decision making. As we hypothesized, our data suggest that lower levels of physiological arousal allowed participants to engage in more strategic thinking, while higher levels of arousal inhibited the deliberative process.

Two statistical considerations also merit attention. First, there could be confounding pre-treatment variables that affect both our measure of emotion (skin

https://doi.org/10.1017/S0020818316000473 Published online by Cambridge University Press

conductance) and outcome decision (whether to choose the war lottery). Controlling for additional pre-treatment covariates of Player B did not change our results.[73] Depending upon how a potential confounder affected our emotional measure and outcome, our results could be biased down or up. Unfortunately, the formal sensitivity analysis proposed by Imai, Keele, and Tingley that deals with omitted variables affecting both the mediator and outcome is not possible here because our treatment variable is continuous rather than binary.[74] Second, we note that our measures of emotional arousal will come with some degree of measurement error. Given that the direct and indirect effects go in opposite directions, if anything we are likely to be underestimating the role of the emotional pathway.[75]

Follow-up Experiment: Shifting Power but No Shifting Offers

In examining the effects of commitment problems by contrasting two experimental conditions, one with a large shift in power and one with a shift so small as to be negligible, we found that the latter condition did not generate a commitment problem by virtue of the fact that the shift in power was too small. This is one way to examine commitment problem dynamics, but it is by no means perfect. For example, the *BigShift* condition might have primed respondents to think about what a "fair" reference point would be.[76] An alternative manipulation of the commitment problem is to allow a large shift in power but prevent offers from changing between Periods 1 and 2.[77] If a reduction in concerns about shifting power occurs, then we should expect that the mediating role of skin conductance should also decrease.[78]

Figure 5 presents the results from a follow-up study similar to our *BigShift* condition from Study 1, but with one important difference: offers are not allowed to change between periods (whatever offer is made in Period 1 is automatically made in Period 2). Study 2 used thirty-six new subjects from the lab's subject pool. The results show that while increases in offer size lead to decreases in the probability that Player B chooses a war lottery, the mediating role of our physiological measure is now no longer statistically different from zero, with the point estimate quite close to zero.

In other words, once we eliminated the commitment problem through the imposition of institutional changes—here changing the rules of the game to prevent shifts in offers—the affective mechanism linking offer size to rejections disappears,

73. We included a measure of right-wing authoritarianism, generalized trust, and ideology.

74. Imai, Keele, and Tingley 2010. Further, we note that while offers are not randomized within a condition, they are not a choice variable of the person who is responding to the offer.

75. VanderWeele, Valeri, and Ogburn 2012.

76. Kahneman 1992.

77. Quek 2017.

78. An alternative design would be to introduce communication where players could try to make verbal commitments. This may or may not have the desired effect depending on whether cheap talk can be persuasive. Tingley and Walter 2011a.

https://doi.org/10.1017/S0020818316000473 Published online by Cambridge University Press

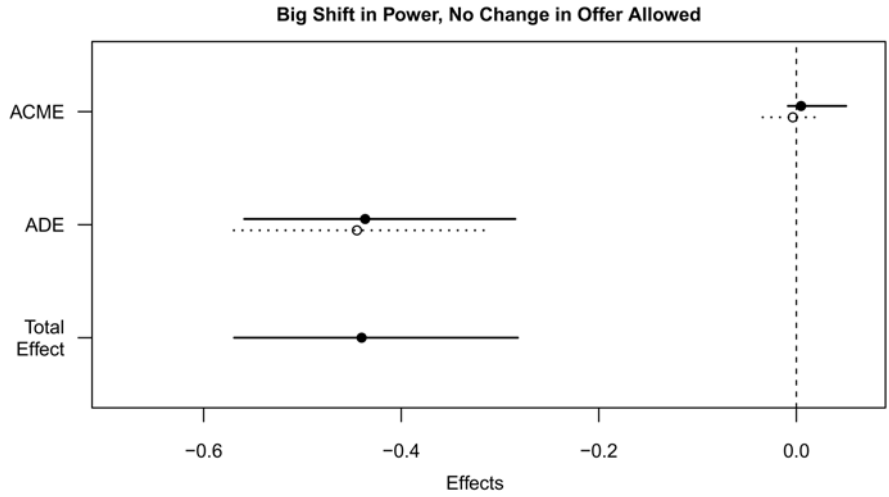


FIGURE 5. *Mediation effects for offer change of 5 to 10 using experimental condition that allows for large shift in power, but no ability for offers to change from round 1 to round 2, with 95% confidence intervals allowing for clustering at the individual level*

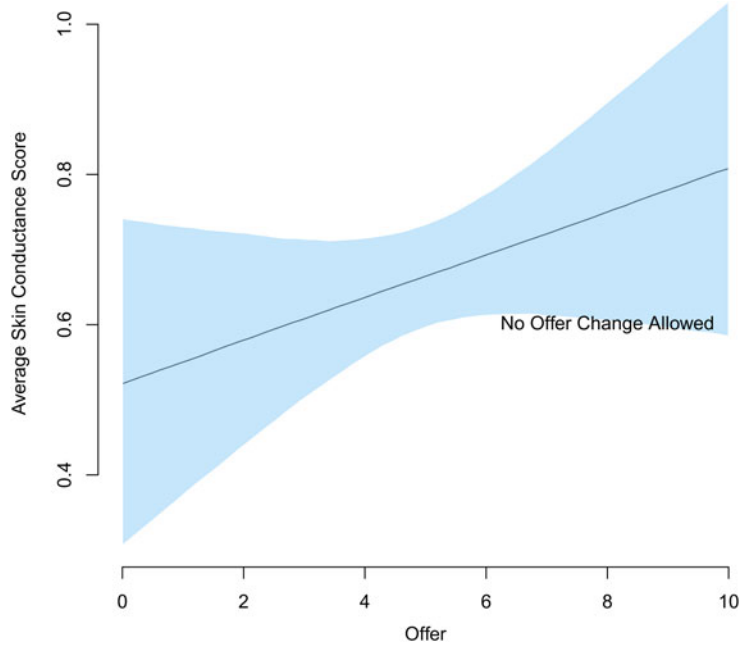


FIGURE 6. *Effect of offer sizes on skin conductance when large power shift and no opportunity for changes in offer 95% confidence intervals with robust standard errors*

https://doi.org/10.1017/S0020818316000473 Published online by Cambridge University Press

consistent with our H4. Thus we have tried to "shut off" the commitment problem in two ways, with both having the same consequences for our physiological measures. In Figure 6 we present the simulated impact of offers on our skin conductance measure, and we see the same pattern as with the *SmallShift* condition in Figure 3. By including this additional design we can be more confident that our results relate to concerns about changes in offers that would occur in the future as a result of shifting power.

Discussion and Conclusion

Explanations for the failure of bargaining and the conflict that results have centered heavily in recent years on so-called "commitment problems" that arise from shifting power. This set of explanations is compelling, but makes no attempt to connect to what individual, human, decision makers might experience. At the same time, a growing body of work in IR has made a strong case for the importance of emotions in political decision making, though with an emphasis on theoretical frameworks and single-case studies rather than systematic evidence. We synthesized these two bodies of literature by proposing a theory of "aversive arousal" that is activated under conditions of large power shifts (but not, it should be noted, under stable conditions). In contrast to canonical models of commitment problems, we argued and found evidence for the notion that physiological and emotional arousal—measured here as skin conductance reactivity—may impair individuals' ability to look forward and think backward to make the strategically optimal choice. In our studies, individuals who knew their power was declining but experienced intense physiological arousal accepted "bad" offers from their counterparts rather than take their chances in conflict, as they should have done (and as individuals who did not experience physiological arousal did).

We also found evidence in support of several critical expectations resulting from the formal model. For example, more subjects preferred to take their chances with war in the present when power was shifting against them than when conditions were stable, a finding that is also consistent with the dual-process model of cognition and System 2 processing. We also found that guaranteeing that offers could not change in the future, mimicking the function of international institutions, shuts off the commitment problem, again providing support for a key pillar of IR theory.

We have brought together a somewhat unusual combination of formal modeling to derive expectations, psychological theories to suggest additional hypotheses, and experimental methods and psychophysiological data to test those hypotheses. While this particular combination is not necessary in all cases, we hope that it serves as an example of one promising way to test the micro-foundations of theories of conflict in a laboratory setting. Unlike previous studies, we carefully design our experiment in a way that lets us analyze the mediating role of physiological responses in a bargaining setting using new statistical tools.

https://doi.org/10.1017/S0020818316000473 Published online by Cambridge University Press

Of course, any successful experiment inevitably generates more questions than it answers. In this case, we focus on emotional arousal more broadly and the dual-process model of cognition. However, a large literature on emotions and bargaining suggests that *specific* emotions (for example, anxiety, anger) may be an important component of the process we describe. This dovetails with what is already a focus of the theoretical literature on emotions in IR. Although the measurement of specific emotions is quite a bit more tenuous than generalized arousal, future advances in methodology seem likely given researchers' intense focus on these issues.[79]

Just as important, one might continue to add additional layers of complexity to the bargaining game we used to more closely mimic important aspects of the international system. We began this process by incorporating international institutions that prevented shifts in offers into our follow-up experiment, but other extensions are possible. For example, despite the caricature sometimes painted of elite leaders sitting behind desks making critical decision in isolation, most decisions take place in a group context. In this special issue, Saunders focuses on advisers, finding that elite leaders are differentially able to monitor their advisers, changing the quality and type of advice they receive.[80] Future work might incorporate these to begin to tackle the "aggregation problem" and construct a plausible and comprehensive framework for behavioral international relations.[81]

There remain important questions about the applicability of such experiments to the study of international relations. Our study uses a convenience sample. We think the best way analytically to evaluate the impact of this is to "sign the bias." Recent work makes a good point that student subject pools and elite samples may differ.[82] Crucially, though, it is important to think about whether any of the effects we identified would be magnified or decreased were we to use an alternative sample. Such considerations are beyond our scope here, but our paper provides an innovative platform for such a discussion. Future studies might investigate important questions about selection into crises. Perhaps those people who are more easily aroused also enter into zero-sum situations more readily because they are less able to recognize potential for joint gains. Future experiments, and papers, designed with this in mind would speak to key themes in the study of international negotiation.[83]

79. One possibility is described in Myers and Tingley 2016.
80. Saunders 2017.
81. Hafner-Burton et al. 2017.
82. Hafner-Burton, Hughes, and Victor 2013; Renshon 2015. Although other recent work identifies similarities in the decision making of leaders and the general public. See Loewen et al. 2014 and Renshon, Yarhi-Milo, and Kertzer 2015.
83. Odell and Tingley 2013.

https://doi.org/10.1017/S0020818316000473 Published online by Cambridge University Press

Supplementary Material

Supplementary material for this article is available at https://doi.org/10.1017/S0020818316000473.

References

Allison, Graham T. 1969. Conceptual Models and the Cuban Missile Crisis. *The American Political Science Review* 63 (3):689–718.

Ariely, Dan, Uri Gneezy, George Loewenstein, and Nina Mazar. 2009. Large Stakes and Big Mistakes. *The Review of Economic Studies* 76 (2):451–69.

Bayram, A. Burcu. 2017. Due Deference: Cosmopolitan Social Identity and the Psychology of Legal Obligation. In The Behavioral Revolution and International Relations, supplement, *International Organization* 71 (S1):S137–S163.

Bleiker, Roland, and Emma Hutchison. 2008. Fear No More: Emotions and World Politics. *Review of International Studies* 34 (1):115–35.

Blight, James G. 1992. *The Shattered Crystal Ball: Fear and Learning in the Cuban Missile Crisis*. Lanham, MD: Rowman and Littlefield.

Blount, Sally. 1995. When Social Outcomes Aren't Fair: The Effect of Causal Attributions on Preferences. *Organizational Behavior and Human Decision Processes* 63 (2):131–44.

Camerer, Colin F. 2003. Behavioural Studies of Strategic Thinking in Games. *Trends in Cognitive Sciences* 7 (5):225–31.

———. 2011. *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton, NJ: Princeton University Press.

Civai, Claudia, Corrado Corradi-Dell Acqua, Matthias Gamer, and Raffaella I. Rumiati. 2010. Are Irrational Reactions to Unfairness Truly Emotionally Driven? Dissociated Behavioural and Emotional Responses in the Ultimatum Game Task. *Cognition* 114 (1):89–95.

Crawford, Neta C. 2000. The Passion of World Politics: Propositions on Emotion and Emotional Relationships. *International Security* 24 (4):116–56.

Damasio, Antonio. 2008. *Descartes' Error: Emotion, Reason and the Human Brain*. New York: Random House.

De Sousa, Ronald. 1987. *The Rationality of Emotion*. Cambridge, MA: MIT Press.

Dowty, Alan. 1984. *Middle East Crisis: US Decision Making in 1958, 1970, and 1973*. Berkeley, CA: University of California Press.

Evans, Jonathan. 2008. Dual-Processing Accounts of Reasoning, Judgment, and Social Cognition. *Annual Review of Psychology* 59:255–78.

Evans, Jonathan, and David E. Over. 1996. *Rationality and Reasoning*. Hove, UK: Psychology Press.

Fearon, James. 1995. Rationalist Explanations for War. *International Organization* 49 (3):379–414.

Gardner, Wendi L., Shira Gabriel, and Amanda B. Diekman. 2000. Interpersonal Processes. In *Handbook of Psychophysiology,* 2nd ed., edited by John T. Cacioppo, Louis G. Tassinary, and Gary G. Berntson, 643–64. New York: Cambridge University Press.

Gigerenzer, Gerd. 2007. *Gut Feelings: The Intelligence of the Unconscious*. New York: Penguin.

Gigerenzer, Gerd, and Terry Regier. 1996. How Do We Tell an Association From a Rule? Comment on Sloman (1996). *Psychological Bulletin* 119 (1):23–26.

Gilinsky, Alberta S., and Ben B. Judd. 1994. Working Memory and Bias in Reasoning Across the Life Span. *Psychology and Aging* 9 (3):356–71.

Grimm, Veronika, and Frederike Mengel. 2011. Let Me Sleep on It: Delay Reduces Rejection Rates in Ultimatum Games. *Economics Letters* 111 (2):113–15.

https://doi.org/10.1017/S0020818316000473 Published online by Cambridge University Press

Hafner-Burton, Emilie M., Stephan Haggard, David A. Lake, and David G. Victor. 2017. The Behavioral Revolution and International Relations. In The Behavioral Revolution and International Relations, supplement, *International Organization* 71 (S1):S1–S31.

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.

Haidt, Jonathan. 2001. The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgment. *Psychological Review* 108 (4):814–34.

———. 2012. *The Righteous Mind: Why Good People Are Divided by Politics and Religion*. New York: Random House.

Hall, Todd H., and Andrew A.G. Ross. 2015. Affective Politics After 9/11. *International Organization* 69 (4):847–79.

Harris, David. 2004. *The Crisis: The President, the Prophet, and the Shah—1979 and the Coming of Militant Islam*. New York: Little, Brown, and Company.

Herrmann, Richard K. 2017. How Attachments to the Nation Shape Beliefs About the World: A Theory of Motivated Reasoning. In The Behavioral Revolution and International Relations, supplement, *International Organization* 71 (S1):S61–S84.

Holsti, Ole R. 1970. Individual Differences in "Definition of the Situation." *Journal of Conflict Resolution* 14 (3):303–10.

Holt, Charles A., and Susan K. Laury. 2001. Varying the Scale of Financial Incentives Under Real and Hypothetical Conditions. *Behavioral and Brain Sciences* 24 (3):417–17.

Imai, Kosuke, Luke Keele, and Dustin Tingley. 2010. A General Approach to Causal Mediation Analysis. *Psychological Methods* 15 (4):309–34.

Kahneman, Daniel. 1992. Reference Points, Anchors, Norms, and Mixed Feelings. *Organizational Behavior and Human Decision Processes* 51 (2):296–312.

———. 2003. A Perspective on Judgment and Choice: Mapping Bounded Rationality. *American Psychologist* 58 (9):697–720.

———. 2011. *Thinking, Fast and Slow*. New York: Macmillan.

Kahneman, Daniel, and Amos Tversky. 1979. Prospect Theory: An Analysis of Decision Under Risk. *Econometrica* 47 (2):263–91.

Kertzer, Joshua D. 2017. Resolve, Time, and Risk. In The Behavioral Revolution and International Relations, supplement, *International Organization* 71 (S1):S109–S136.

Khong, Yuen F. 1992. *Analogies at War: Korea, Munich, Dien Bien Phu, and the Vietnam Decisions of 1965*. Princeton, NJ: Princeton University Press.

Lambourne, Kate, and Phillip Tomporowski. 2010. The Effect of Exercise-Induced Arousal on Cognitive Task Performance: A Meta-Regression Analysis. *Brain Research* 1341:12–24.

Levitt, Steven D., and John A. List. 2007. What Do Laboratory Experiments Measuring Social Preferences Reveal About the Real World? *The Journal of Economic Perspectives* 21 (2):153–74.

Loewen, Peter John, Lior Sheffer, Stuart N. Soroka, Stefaan Walgrave, and Tamir Shaefer. 2014. Are Politicians Better Risky Decision Makers? Evidence from Experiments with Members of Parliament in Belgium, Canada, and Israel. Working paper.

Loewenstein, George, and Jennifer S. Lerner. 2003. The Role of Affect in Decision Making. In *Handbook of Affective Sciences*, edited by Richard J. Davidson, Klaus R. Scherer, and H. Hill Goldsmith, 619–42. New York: Oxford University Press.

McDermott, Rose. 2004a. The Feeling of Rationality: The Meaning of Neuroscientific Advances for Political Science. *Perspectives on Politics* 2 (4):691–706.

———. 2004b. *Political Psychology in International Relations*. Ann Arbor, MI: University of Michigan Press.

Miller, Patrick R. 2011. The Emotional Citizen: Emotion as a Function of Political Sophistication. *Political Psychology* 32 (4):575–600.

Mutz, Diana C. 2007. Effects of "In-Your-Face" Television Discourse on Perceptions of a Legitimate Opposition. *American Political Science Review* 101 (4):621–35.

https://doi.org/10.1017/S0020818316000473 Published online by Cambridge University Press

Myers, Dan, and Dustin H. Tingley. 2016. Influence of Emotions on Trust. *Political Analysis* 24 (4):492–500.

Odell, John, and Dustin Tingley. 2013. Negotiating Agreements in International Relations. International Relations Working Group of the American Political Science Association 2013 Task Force on Negotiating Agreements in Politics. Available at: <http://www.apsanet.org/portals/54/Files/Task%20Force%20Reports/Chapter7Mansbridge.pdf>. Accessed 30 January 2017.

Oneal, John R. 1988. The Rationality of Decision Making During International Crises. *Polity* 20 (4):598–622.

Osman, Magda. 2004. An Evaluation of Dual-Process Theories of Reasoning. *Psychonomic Bulletin and Review* 11 (6):988–1010.

Ostrom, Elinor. 1998. A Behavioral Approach to the Rational Choice Theory of Collective Action: Presidential Address, American Political Science Association, 1997. *American Political Science Review* 92 (1):1–22.

Powell, Robert. 2006. War As a Commitment Problem. *International Organization* 60 (1):169–203.

Quek, Kai. 2017. Rationalist Experiments on War. *Political Science Research and Methods* 5 (1):123–42.

Rathbun, Brian C., Joshua D. Kertzer, and Mark Paradis. 2017. *Homo Diplomaticus*: Mixed-Method Evidence of Variation in Strategic Rationality. In The Behavioral Revolution and International Relations, supplement, *International Organization* 71 (S1):S33–S59.

Renshon, Jonathan. 2015. Losing Face and Sinking Costs: Experimental Evidence on the Judgment of Political and Military Leaders. *International Organization* 69 (3):659–95.

Renshon, Jonathan, Julia J. Lee, and Dustin Tingley. 2015. Physiological Arousal and Political Beliefs. *Political Psychology* 36 (5):569–85.

Renshon, Jonathan, and Jennifer Lerner. 2012. The Role of Emotions in Foreign Policy Decision Making. In *Encyclopedia of Peace Psychology*, edited by Daniel J. Christie and Cristina Montiel, 313–17. Malden, MA: Wiley-Blackwell Press.

Renshon, Jonathan, Keren Yarhi-Milo, and Joshua D. Kertzer. 2015. Democratic Leaders, Resolve and War: Experimental Evidence from the Knesset. Working Paper.

Rho, Sungmin, and Michael Tomz. 2017. Why Don't Trade Preferences Reflect Economic Self-Interest? In The Behavioral Revolution and International Relations, supplement, *International Organization* 71 (S1): S85–S108.

Sanfey, Alan G., James K. Rilling, Jessica A. Aronson, Leigh E. Nystrom, and Jonathan D. Cohen. 2003. The Neural Basis of Economic Decision-Making in the Ultimatum Game. *Science* 300 (5626):1755.

Saunders, Elizabeth N. 2017. No Substitute for Experience: Presidents, Advisers, and Information in Group Decision Making. In The Behavioral Revolution and International Relations, supplement, *International Organization* 71 (S1):S219–S247.

Sequeira, Henrique, and Jean-Claude Roy. 1997. Neural Control of Electrodermal Activity. In *Central Nervous Control of Autonomic Function,* Vol. 11, edited by David Jordan, 259–93. Amsterdam: Harwood Academic Publishers.

Sloman, Steven A. 1996. The Empirical Case for Two Systems of Reasoning. *Psychological Bulletin* 119 (1):3–22.

Stern, Sheldon M. 2005. *The Week the World Stood Still: Inside the Secret Cuban Missile Crisis*. Stanford, CA: Stanford University Press.

Tingley, Dustin. 2011. The Dark Side of the Future: An Experimental Test of Commitment Problems in Bargaining. *International Studies Quarterly* 55 (2):521–44.

———. 2017. Rising Power on the Mind. In The Behavioral Revolution and International Relations, supplement, *International Organization* 71 (S1):S165–S188.

Tingley, Dustin, 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.

VanderWeele, Tyler J., Linda Valeri, and Elizabeth L. Ogburn. 2012. The Role of Measurement Error and Misclassification in Mediation Analysis. *Epidemiology* 23 (4):561–64.

https://doi.org/10.1017/S0020818316000473 Published online by Cambridge University Press

Van't Wout, Mascha, Rene S. Kahn, Alan G. Sanfey, and Andre Aleman. 2006. Affective State and Decision-Making in the Ultimatum Game. *Experimental Brain Research* 169 (4):564–68.

Welch, David. 2003. Culture and Emotion as Obstacles to Good Judgment. In *Good Judgment in Foreign Policy: Theory and Application*, edited by Stanley Renshon and Deborah Welch Larson, 483–507. Lanham, MD: Rowman and Littlefield.

Woodward, Bob. 2002. *Bush at War*. New York: Simon and Schuster.

https://doi.org/10.1017/S0020818316000473 Published online by Cambridge University Press