Physiological Responses to Shifting Bargaining Power: Micro-Foundations of Commitment Problems in International Politics[1]

Dustin Tingley[2] Jooa Julia Lee[3] Jonathan Renshon[4]

This draft: November 5, 2012

[1]We are especially grateful to Ariya Hagh, and the team at the Harvard Decision Science Laboratory for facilitating this research, including Judith Ezike, Dan Zangri, Nicole Ludmir, Gabe Mansur, Jeanie Nguyen, Jakob Schneider, Phil Esterman, 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 Paul Poast, Alex Peysakhovich, and Yunkyu Sohn for their thoughtful comments on the previous version of this paper.

[2]Government Department, Harvard University, Cambridge, MA 02138. Email: dtingley@gov.harvard.edu, URL: scholar.harvard.edu/dtingley

[3]PhD Candidate, Harvard Kennedy School. Email: jooajulialee@fas.harvard.edu, URL: http://scholar.harvard.edu/jooajulialee

[4]Assistant Professor, Department of Political Science, University of Wisconsin-Madison. Email: jrenshon@fas.harvard.edu.edu, URL: http://scholar.harvard.edu/renshon

Abstract

Commitment problems are often regarded as one of the most important, and common, causes of international conflict. Instability of bargaining power can generate an incentive to reject offers, and thus lead to a costly conflict that destroys resources. Despite this, there is a noticeable gap in our understanding of how commitment problems change individual behavior in actual bargaining. We use a novel experimental protocol to study how individuals react to the changes in bargaining power in the game, where bargaining power is operationalized as the probability of winning all the resources. We use a psycho-physiological measure – namely skin conductance levels –in order to assess individuals' physiological reactions to bargaining. We find that i) people are more likely to reject offers when they expect the opponent's increasing bargaining power. We also find, only in the presence of shifting power, that ii) in the presence of shifting power, higher offers lead to less arousal, and iii) higher physiological arousal decreases rejection decisions, perhaps due to an increase in 'aversive' physiological arousal. Our novel findings suggest that physiological arousal does not always lead to rejection of unfair offers. Rather, when individuals face large power shifts, arousal may allow more intuitive decision-making and less backward induction. This suggests that under certain circumstances, emotionally-aroused individuals are less prone to commitment problems.

1 Introduction

Given that conflict is costly and destroys resources, why does it occur? A number of scholars have asked this question and proposed various explanations. Rationalist explanations frequently turn on the presence of a commitment problem or incomplete information. One mechanism in particular –commitment problems –has had enormous influence on the study of international conflict and is often argued to be fundamental to understanding the breakdown of peace.

Commitment problems arise because shifts in bargaining strength generate incentives in the future to renege on current commitments. For example, if country i is increasing in strength relative to country j, country i will find it difficult to 'credibly commit' to not taking advantage of j in the future. With this in mind, the actor whose power is declining prefers conflict while they have a bargaining advantage rather than a peaceful resolution that may leave them open to predation in the future. This, understandably, can create conflict despite the fact that conflict is costly (Fearon, 1995, 2004; Powell, 2006). Furthermore, due to liquidity constraints, bargainers cannot prevent their opponents from being concerned about such adverse shifts.[1] Thus, future changes in bargaining power have been used to explain preventive war, whereby a declining state chooses to wage a war against the opponent to take advantage of its bargaining advantages (Fearon, 1995).

In recent years the international relations literature has seen a proliferation of work studying the behavioral foundations of bargaining, drawing heavily on experimental methods that have human subjects bargaining over resources or responding to hypothetical crisis simulations (e.g., Tingley, 2011; Butler et al., 2007; Tingley and Walter, 2011b,a; Tingley and Wang, 2010; Tomz, 2007). For example, Tingley (2011) and McBride and Skaperdas (2012) examine how changes in the likelihood of future bargaining–the "shadow

[1] 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). Elsewhere I study this type of explanation for conflict using experimental methods (Tingley and Wang, 2010). The role of commitment problems in other settings is also well developed theoretically and empirically with observational data (Greif et al., 1994; Simmons, 1994; Simmons and Danner, 2010)

of the future"–influence the extent to which commitment problems create conflict. Some recent studies have begun to measure relevant biological variables in an experimental bargaining environment as well (McDermott et al., 2009; Rosen, 2005; McDermott et al., 2007). While many of these experiments do not explicitly study commitment problems, the ubiquitous role commitment problem explanations play in the international relations literature and relative paucity of empirical research suggests the need for sustained study.

Although commitment problems have generally been studied as part of the rationalist paradigm, we approach the problem from a different perspective by focusing on the role of emotions and intuition. Emotional reactions play an important role in guiding our behavior in strategic situations, and thus are thus a prime candidate to affect how individuals respond to anticipated power shifts. Drawing on a body of research that has identified a critical link between emotionally activated processes and intuitive decision-making (Epstein et al., 1996; Dane and Pratt, 2007), we investigate both affective and cognitive mechanisms that drive individual behavior in bargaining. By closely tracking the level of emotional arousal, we also indirectly measure the extent to which subjects engage in intuitive (versus deliberative) decision-making decision-making. This is accomplished via the measurement of skin conductance reactivity, a physiological measure of arousal in the autonomic nervous system. This allows us to measure participants' sympathetic nervous system activation free from any potentially disruptive social-masking or impression-management effects and in a more direct manner than survey-based measures of emotion.

Using this novel method, we ask two broad questions in this paper. First, how does the size of the power shift (large or small) affect how *individual* decision makers respond? Second, what is the physiological mechanism which governs the reaction of individuals to unfair offers when a power shift is imminent? More specifically, do unfair offers trigger a certain type of processing style, one that is intuitive and affect-laden and more likely to lead to worse outcomes in a bargaining setting? In order to test these, we extend a recently-developed experimental protocol (Tingley, 2011), that allows us to measure the impact of shifting power – large and small power shifts, operationalized by the extend to which the power each player holds differs –in a two-stage bargaining game.

We find that larger power shifts anticipated do in fact lead to higher probabilities of offer rejections. This fits very well with the standard explanations of preventive military action in which declining powers prefer to take their chances *now* rather than face a steep decline relative to their adversaries. However, controlling for offer size, we find that increased emotional arousal (higher skin conductance) led to a lower level of offer rejection. We interpret this finding in light of the dual-process model of cognition, which suggests that 'aversive arousal' is tampering with the ability of those bargaining to act in a strategically-optimal manner. This particular finding suggests that emotional arousal does not necessarily lead to less optimal, inefficient outcome, but may actually run counter to decisions to wage a war.

The paper is organized as follows. First, we provide a model of commitment problems in Section 2 and the theories of physiological arousal on bargaining in Section 3. Section 5 outlines the design of the experiment, while Section 6 outlines our main findings related to physiological arousal and bargaining. Section 7 concludes.

2 Model of Commitment Problems

To see how power shifts can lead to conflict we provide a brief sketch of a simple two period bargaining model.[2] The basic intuition is that when one player expects to be disadvantaged in the future, they have an incentive to start a conflict in order to forestall the shift in bargaining power. There are two players, A and B. In each period there is a resource with value 1. This resource may be divided between the players. The bargaining protocol is as follows. In the first period player A makes a demand, x_{1A}, to player B. This leaves player B with $1 - x_{1A}$. If the demand is accepted, then in the second period there is another bargaining stage, where player A makes a demand x_{2A}. If the first period demand is rejected then both players play a war lottery, with probability p_1 player A wins. One player wins the resource for period 1, but both players pay costs $c_A = c_B$. The winner of the lottery obtains the entire resource in period 2. If the offer is rejected in the second

[2]This is simply a two period version of the model in Fearon (1995).

period, both players pay the cost and play a war lottery, with probability p_2 player A wins.

To analyze the complete information game we use backwards induction. We begin with the second period decision by actor B. In Period 2 player B is indifferent between demand x_{2A} and lottery if $(1-p_2)\times 1 - c_B = 1 - x_{2A}$. Thus if $x_{2A} = p_2 + c_B$ the actor B is indifferent and by assumption accepts the demand. In period 2 player A's expected utility from the lottery is $p_2 - c_A$. Hence A will prefer to make a demand $x_{2A} = p_2 + c_B$ if $p_2 + c_B > p_2 - c_A$. This holds because $c_B = c_A$. They can't make more than this demand because then it will be rejected. So, the optimal demand in the second period is $x^*_{2A} = p_2 + c_B$.

Now consider period 1. B's utility from rejecting the lottery in period 1 is $(1-p_1) - c_B + \delta(1-p_1)$, where δ represents the discount rate for the second period, or the likelihood that the second period is played. In this model we assume that $\delta = 1$. Player B's utility from accepting a demand x_{1A} is equal to $1 - x_{1A} + \delta(1 - p_2 - c_B)$.

Thus player B will reject the demand x_{1A} if

$$\begin{aligned}(1-p_1) - c_B + \delta(1-p_1) &> 1 - x_{1A} + \delta(1-p_2-c_B)\\ x_{1A} &> p_1 + \delta p_1 - \delta p_2 + c_B - \delta c_B\end{aligned}$$

Hence they will be indifferent if $x_{1A} = p_1 + \delta p_1 - \delta p_2 + c_B - \delta c_B$,

Now consider A's expected utility in period 1. If they have the lottery rejected they get $p_1 - c_A + \delta p_1$. If they have some demand x_{1A} accepted then they get $x_{1A} + \delta(p_2 + c_B)$. They will want their first period demand accepted if

$$\begin{aligned}x_{1A} + \delta(p_2 + c_B) &\geq p_1 - c_A + \delta p_1\\ x_{1A} &\geq p_1 + \delta p_1 - \delta p_2 - c_A - \delta c_B\end{aligned}$$

Now note that the RHS of this is almost identical to what will make B indifferent, except that it is slightly smaller ($-c_A$ instead of c_B). Hence A will make demand $x^*_{1A} = p_1 + \delta p_1 - \delta p_2 + c_B - \delta c_B$ and $x^*_{2A} = p_2 + c_B$. The key question is when the first period demand will be rejected. This will occur when x^*_{1A} is less than 0. That is, when player A can not offer (demand little) enough to make player B accept the demand in light of what they expect to get in period 2. This holds when we have $0 > p_1 + \delta p_1 - \delta p_2 + c_B - \delta c_B$. Assuming $\delta = 1$ this reduces to $p_2 - 2p_1 > 0$. Thus as p_2 gets larger and/or p_1 gets smaller, this condition is more likely to hold.

In the experiment that follows, our condition where we expect preventive war to occur has two different power levels, $p_1 = .3$ and $p_2 = .7$ and our condition where we expect preventive war to not occur has $p_1 = .49$ and $p_2 = .51$. In the first condition there is a large difference between p_1 and p_2 whereas in the second condition there is not. As in the model, we operationalize these values of power as the probability one wins the entire resource following a rejected offer.

3 Physiological Arousal and Bargaining

The theoretical models in this literature identify "mechanisms" that lead to conflict in the sense that a set of utility maximization assumptions and a particular bargaining environment result in particular predictions about when conflict will occur. However, they do not offer mechanisms particular to the human decision-makers that are ultimately of interest. 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 simply relying on the abstract choice models themselves (for brief review see Camerer 2003a, chapter 2). 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? Parallel with other literatures in political science using a novel method in psychophysiology (e.g., Oxley et al., 2008), we study the physiological bases of bargaining in situations where commitment problems are present. Before presenting our research design and measurement strategy, we build on previous work to develop several hypotheses about how individual physiological responses and bargaining outcomes.

3.1 Intuitive vs. Reflective Bargaining

The idea that individuals' cognitive processing can be described as two parallel systems can be traced back at least as far as James (1890) and Freud (1900), who described a dual system of "rational" and "irrational' thinking.[3] We follow Evans and Over (1996) in conceptualizing two parallel cognitive processes, System I and System II. The former is based upon previous experiences and beliefs. It is characterized as 'implicit, associative, fast and highly robust (Gilinsky and Judd, 1994).' Perhaps most importantly, it is automatic: it occurs all the time without our awareness. In contrast, System II is slow, controllable and explicit (in the sense that one is consciously engaging in the act of analytical reasoning). Emotions tend to be more closely associated with System I processing. That is, individuals seem to feel an emotion, and then construct a rational edifice around that feeling in order to explain why they are angry/sad/happy/etc. (Haidt, 2001). In some cases, emotions are explicitly linked to System I processing (Epstein, 1994; Hassin et al., 2004; Evans, 2008), while in others evidence is provided based on the parts of the brain involved in both emotion-processing and System I processes (Lieberman, 2003).

While there is little work (of which we are aware) which applies the dual-process framework directly to bargaining, there is a burgeoning literature on the role of emotions in experimental economics. These behavioral investigations of bargaining have typically focused on the "ultimatum game (UG)." In the UG one person proposes a division of a resource to a second person. If the second person accepts, then the resources 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. Or, in other words, the standard equilibrium analysis of the UG is that there should never be rejected offers, even when they are arbitrarily low. While the specific behavioral predictions from our game differ in significant ways (see Section 2), previous work on cognitive processing in the UG is still relevant to the current study.

Initial evidence from the UG suggests that emotions play a powerful role, despite

[3]For more detailed history of dual-process frameworks, see Osman (2004). For an overview of some of the controversy associated with the framework, see Gigerenzer and Regier (1996).

assumptions that strategic thinking –"cold calculation" –should dominate. Blount (1995) used a variant of the UG to show that unfair offers by the proposers can evoke anger in responders when the allocation was intentional, but not when it was randomly generated. Pillutla and Murnighan (1996) also found that responders who have full information about the total endowment available to the proposer rejected more than those who did not have any information. In other words, emotions like anger can be triggered under certain circumstances, but there is a logic to how this process occurs; a "bad" offer by itself is not enough to trigger anger. There must be some knowledge of the motivation behind the allocation (to know whether there was an *intent* to be unfair) or a larger understanding of the amount available to the proposer (to know if the allocation was fair).

Advances in neurophysiology have supported these findings. Sanfey et al. (2003) used functional neuroimaging to record the brain activity of responders who were faced with different distribution of allocations. They observed that unfair offers resulted in a higher level of activation in bilateral anterior insula, an area of the brain associated with negative emotional arousal. The intensity of activation in this region was significantly correlated with the extent to which the offers were unfair, and thus led to a higher probability of rejection. Van't Wout et al. (2006) and Civai et al. (2010) similarly observed increased electrodermal activity (skin conductance), which is associated with insula activity, when responders received unfair offers. The relationship between offer size, physiological arousal, and probability of rejection was consistent across these studies. Drawing from this research, we hypothesize that unfair offers in our bargaining game (despite its difference set-up from the traditional UG) should trigger greater physiological arousal than fairer offers.

However, the link between unfairness, physiological arousal and the act of rejection is less clear in the current study, which differs from the traditional UG in significant ways. In the UG, rejection is punitive; responders are hurting themselves by rejecting low/unfair offers, but they are also hurting the proposer, who gets nothing. One possible explanation for this behavior is that unfair offers in the UG play upon our aversion to inequity or unfairness, which triggers and and the resulting choice to reject (Yamagishi et al., 2009).

Our game involves two stages, in contrast to the traditional ultimatum game set-up in which actors face a "one-shot" choice. Unlike the standard ultimatum game (in which rejection leads to both players receiving 0), the responder's rejection leads to a gamble where one of the players recieves all of the resources. In the *declining power* condition, the responder wins all of the resources with a probability of .7 in Round 1 and .3 in Round 2 (the probabilities are reversed for the proposer, who wins with a probability of .3 in Round 1 and .7 in Round 2). In the *control* condition, the probability of winning stays at roughly .5 for both rounds. In other words, in the *declining power* condition, both parties know that the proposer will be gaining power in the second round, 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 Round 1, when they still enjoy a power advantage.

Given these notable differences, we hypothesize that these factors may change the way unfair offers and their physiological responses affect rejection rates. This might happen for several reasons. First, the relationship between unfair offers and rejection may not hold true if a responder cannot effectively punish the unfair proposer. For example, Yamagishi et al. (2009) conducted a variant of the ultimatum game, where rejection did not influence the monetary payoff of the proposer. Although they still found that 33-34% of participants rejected the clearly unfair offers (those smaller than 20-30%), potentially driven by anger and moral disgust, this rejection rate was much lower than those found in the standard ultimatum game (approximately 50%). This suggests that, despite the emotional urge to punish the unfair proposer, the absence of effective punishment in this game deterred some participants from acting on their impulses. Similarly, in our bargaining setting, rejection does not necessarily punish the unfair proposer, since both players face uncertain outcomes.

Second, physiological arousal may not necessarily lead to high rejection rates in all circumstances. Instead, we focus on the possibility that physiological arousal may tamper with a more deliberate, cognitive processes of strategic thinking. We draw on the dual-process framework (Kahneman, 2003; Sloman, 1996) that decisions are made by the interaction of intuitive and reflective processes. Physiological arousal may give rise to more intuitive decision-making, which is associated with fast, automatic, and effortless processing (Lambourne and Tomporowski, 2010). On the other hand, the absence of physiological arousal would allow more reflective decision-making, which is associated with slow, effortful, and deliberate processing. A good example of both intuitive and reflective processes in the ultimatum game can be found in the work of Grimm and Mengel (2011), who found that time delays allowed more deliberation, and thus decreased rejection rates.

However, in our bargaining game, the presence of large power shifts implies that the responder's dominant strategy is reject any offers in Round 1. Put simply, the responder's chances in Round 1 will always be better than in Round 2 under conditions of declining power. Choosing this strategy requires the ability to look ahead and induct backwards, but as Camerer (2003b) argues, people's ability to engage in such strategic, iterated thinking is limited. We thus predicted that responders with lower physiological arousal should be more likely to reject Round 1 offers since their ability to engage in reflective processing would not be impaired (compared to those with higher physiological arousal). 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 Round 2.

Taken together, we hypothesize that the size of offers will predict rejection rates, and this relationship will be explained by the responder's physiological arousal, measured by skin conductance levels. More specifically, we expect that skin conductance reactivity will mediate the relationship between high offers and higher rejection rates.

4 Hypotheses

Based on the previous discussion we have several hypotheses. First is the basic hypothesis from the game theoretic model that when there is a larger power shift then we should see more rejected offers. We operationalize this by investigating whether for any given offer there will be a higher probability it is rejected when there is a large power shift (Tingley, 2011).

Hypothesis 1: When there is a large power shift, for any proposed division (offer size), the probability of rejection will be higher.

The second hypothesis concerns the effect of offers on skin conductance levels. Following previous work, we expect that low offers will (because they are likely to be perceived as unfair) lead to higher levels of physiological arousal, but only in the presence of a large power shift. This follows from the logic that the presence of a shifting power dynamic is likely to make unfair offers more psychologically salient, and aversive, than in conditions where conditions remain stable across periods. In other words, the proposer's unfair offer in Round 1 would 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 may not activate the same aversive physiological responses against the low offer, as the responder is not as likely to be concerned with future bargaining, but may be more concerned with the cost of rejection.

Hypothesis 2: When there is a large power shift, low offers should be associated with higher skin conductance levels. This relationship is likely to be attenuated or muted under conditions in which power conditions remain stable across periods.

Our final hypothesis concerns the influence of arousal on rejection decisions, controlling for offer size. In other words, what is the effect of offers on rejection rates that are transmitted though changes in arousal, when the players are expecting a large power shift? Following the dual process framework, we predict that the effect of low offers on rejection rates that is transmitted through though greater physiological arousal will be positive; that is, it will lower rejection rates. The effect of high offers will lead to higher rejection rates, mediated by lower levels of physiological arousal.

Hypothesis 3: 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.

5 Experimental Design

5.1 Lab Procedure

We conducted this study at the Harvard Decision Science Laboratory using 121 male undergraduate students (mean age 22) signed up. Each session began with subjects being hooked up to the physiological sensors described below. 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 level[4] the bargaining games began, described in Section 5.2. Each subject was randomly paired with another subject and randomly assigned a position as described below in more detail. 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, through parallel port communication. This enables us to precisely match a given subject's physiological response at a given time to specific actions in the bargaining game. After playing 10 matches, subjects took a short post-experiment 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 50 minutes of time.

[4]Subjects watched a video (2m49s) featuring relaxing images of beaches and palm trees with calm music in the background to measure the physiological responses at the baseline.

5.2 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 the second period each subject was randomly matched with another subject and played the same 2 period game. Each period of the game had the same basic structure. In period 1, player A proposes a division of a resource of size $R = 10$, $(x_{1A}, x_{1B} = 10 - x_{1A})$. Here the subscripts represents 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, the entire resource is assigned by the computer to player A or B with probability p_1 that A wins the resource and $1 - p_1$ B wins, where the subscript again represents period 1. Both players also pay a cost $c_A, c_B = 2$. Following rejection in the first period, in the second period the winner of the resource receives $R = 10$ points automatically. If the first period offer is accepted, in the second period the same game is played with $R = 10$ points being divided and both players stay in the same position (A or B).

The experiment had two treatment conditions that determined how the probability of gaining the resources shifted from one period to the next. 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) such that in Period 2 the power dynamics have reversed. In this situation, our game-theoretic model predicts that player B is likely to reject offers in Period 1, since 'taking their chances' in a Period 1 gamble is preferable to doing so in Period 2 once a major power shift has occurred. This is illustrative of 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 the second period), 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.

All aspects of the game are common knowledge. Subjects were paid based on one randomly selected repetition of the experiment and received $1 for every 10 points earned.

5.3 Physiological Data Acquisition

Participants' tonic skin conductance levels (SCLs) were measured continuously throughout the experiment. SCL is a reliable indicator of increased attention and arousal, and is associated with greater insula activities (for reviews, see Christie, 1981; Sequeira and Roy, 1997). Skin conductance has also been used in similar contexts in past research to investigate the relationship between actors' emotions (measured by physiological arousal) and their decisions (Van't Wout et al., 2006; Civai et al., 2010). Despite its important advantages, one potential shortcoming of this measure of physiological arousal is that it does not allow us to distinguish between different discrete emotions (i. e., anxiety vs. anger), or the valence of general feelings (i.e., positive or negative feelings). However, given that one's autonomic nervous system responses precede conscious awareness, and are difficult to control consciously (Gardner et al., 2000; Bechara et al., 1997), we believe that the fluctuation (i.e., reactivity) in skin conductance levels should provide valuable information about the extent to which intuitive vs. deliberate processes are in being utilized.

Skin conductance data for all participants was collected through the use of two disposable Biopacs (Santa Barbara, CA) electrodes (Model: EL507), which are filled with Biopac Skin Conductance Electrode Paste. These electrodes were placed on the palms of the participant's nondominant hand (the thenar and hypothenar eminences). SCL was then checked by research assistants to ensure proper recordings, and following that participants watched a short video (2m49s) featuring images of beaches and palm tress with calm music. SCL measurements of this time period were used as 'baseline' physiological arousal.

We followed the standard procedure of scoring the skin conductance data; raw measures were amplified using a gain of $25\mu\Omega$ and a low-pass filter of 5Hz. They were then calculated in microsiemens values using Mindware software EDA module 3.0 (Mindware Technologies, Gahanna, OH) by research assistants who were blind to both the study hypothesis and conditions. We then output the scored data into a time series. Finally, in order to address potential individual differences in variability in skin conductance, our data were transformed to deviations from the baseline condition (using the average SCLs while the video was played) and standardized within each participant (see Ben-Shakhar, 1985; Bush et al., 1993).

Our measure of SCL use this standardized deviation from baseline score. To do this our experimental design requires that we extract these standardized scores during the period where where decisions were made. We were able to do this because of the our procedure of placing tags in the time series via parallel port communication (as discussed above). Then, we took the average SCL deviation from the time *after* the offer was made, and *before* responders make the decision to accept or reject the offer in the first round. This average deviation from baseline during this particular phase of the experiment is our key physiological variable.[5]

6 Results

We begin with simple descriptive statistics of the average first round offer in both conditions. Figure 1 presents histograms of the amount offered to actor B with horizontal lines indicating the average offers in the experimental conditions. Average offers in the BigShift condition were slightly higher. Furthermore, the variation in offers was also much higher in the BigShift condition. While the modal answer in both conditions was a 50/50 split, there were many more offers above and below this point in the SmallShift condition. Next we move to a discussion of our key hypotheses.[6]

[5]Future versions of the paper will also calculate reaction times by subject. Our physiological results are robust to using medians instead of means. Any deviations above 3 are capped at 3 and below -3 is capped at -3, though our results hold not making this common restriction which was rare in the data.

[6]Future versions of the paper could also investigate learning effects.

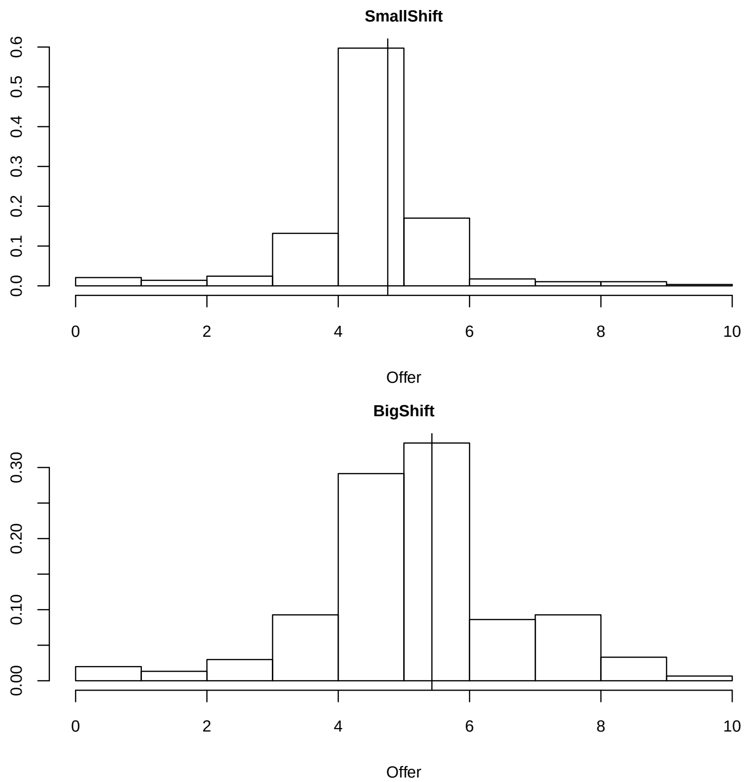

Figure 1: Histogram of first round offers by power shift condition.

6.1 Hypothesis 1: Experimental condition, offers and rejection decisions

Our results were consistent with hypothesis 1, that offers are more likely to be rejected in the BigShift condition. We estimated a probit regression model where the dependent variable is whether or not the offer was rejected. The independent variables were the size of the offer, an indicator for whether the experimental session used the BigShift or SmallShift condition, 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.

Offers below 5 are almost always rejected in both the BigShift and SmallShift conditions. Above an offer size of about 5, offers are more likely to be rejected when there is a large power shift. That is, for more generous offers, individuals in both conditions are more likely to accept the offers than less generous offers. However, for any of these more generous offers they are more likely to be rejected in the BigShift condition.

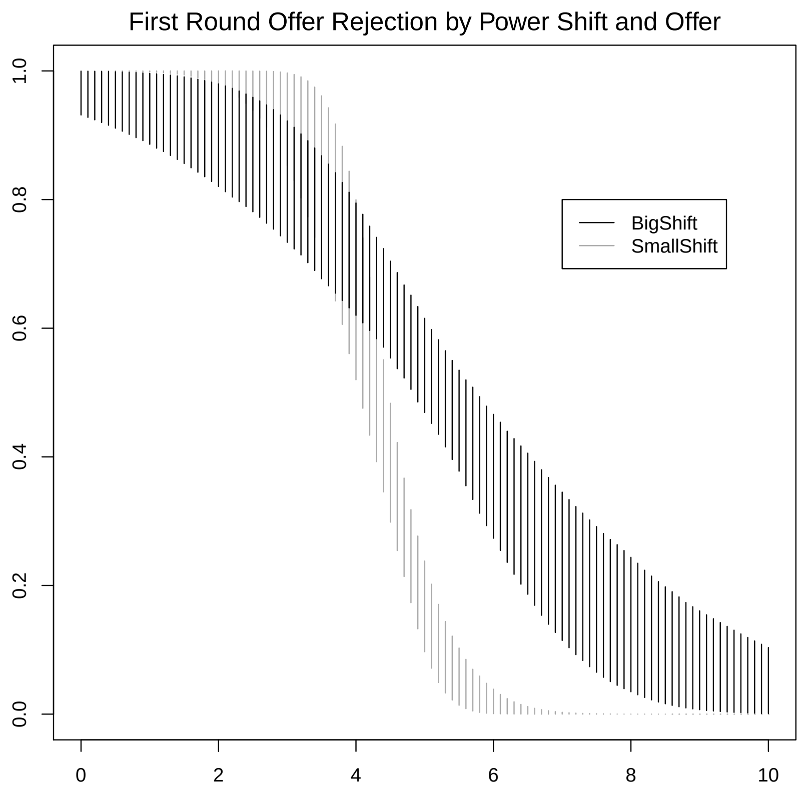

Figure 2: Probability offer rejected as a function of offer size and whether power shift large or small.

6.2 Hypothesis 2: Effect of offers on skin conductance

Next we investigates whether offer effects arousal in the two experimental conditions. To do this we estimated a simple linear regression model with our measure of arousal as the dependent variable and offers as the independent variable.[7] 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 offers and arousal, t=-.3, *n.s.* To display the results Figure 3 plots predicted arousal as a

[7]We also estimated generalized additive models that does not impose a linearity assumption. Results were nearly identical to the linear model.

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. Unfair offers made by the rising power in the first round are perceived more provocative, and may thus be interpreted as a challenge, which might have activated more physiological arousal among the responders. It should also be noted that the relationship between offer size and skin conductance reactivity was more pronounced in the BigShift condition, as compared to the SmallShift condition. This finding indicates that the typical outcomes of low offers – aversive arousal – in the ultimatum game setting may not be directly applied to the bargaining setting. Due to the repeated nature of the bargaining, the anticipation of the power shifts significantly moderates the effect of low offers on physiological arousal.

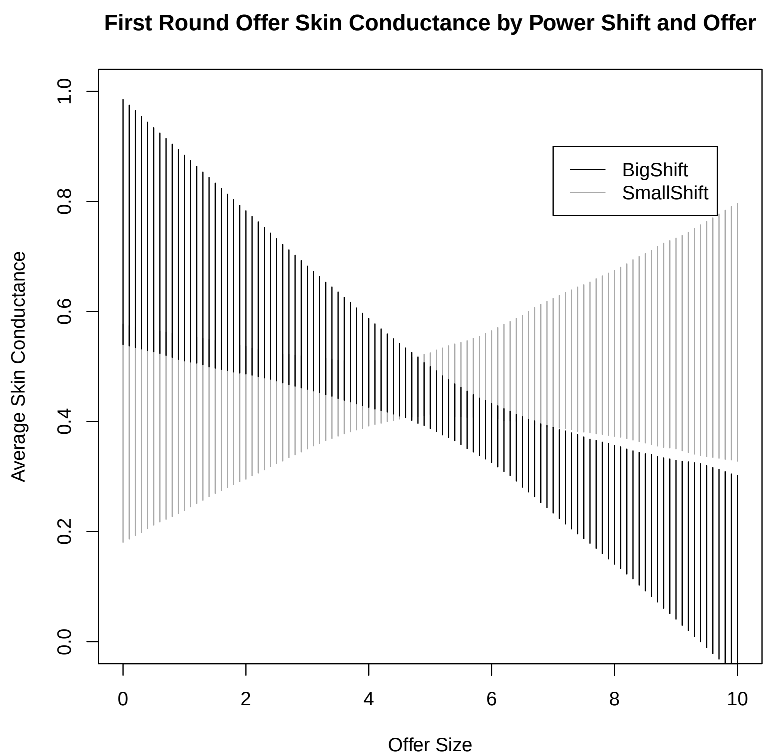

Figure 3: Effect of offer sizes on skin conductance. 80% confidence intervals used for presentational purposes and robust standard errors.

6.3 Hypothesis 3: Mediation of offer effect via arousal

Finally we turn to the role of physiological arousal in mediating the relationship between offers and rejection decisions. The mediation effect is the change in the outcome while holding the treatment condition constant but varying the values of the mediating variable. Formally, the mediation effect under treatment condition t, $\delta_i(t)$ can be defined at the individual level as:

$$\delta_i(t) \quad \equiv \quad Y_i(t, M_i(t)) - Y_i(t, M_i(t')), \tag{1}$$

where $Y_i(t, M_i(t))$ represents the value of the outcome variable under the treatment t. As discussed elsewhere (Imai et al., 2011) inference about this requires a sequential ignorability assumption, where 1) there is no omitted variable that affects for the treatment and outcome and 2) there is no omitted variable that causes both the mediating variable and outcome variable. In the current design, the offer is not randomized. But it is implausible that rejection decisions by a different person could be influenced by the same confounding variable that influences offer decisions.[8] The second assumption is more difficult to deal with and can be approached either through inclusion of pre-treatment controls or sensitivity analysis. We discuss each below.

We estimate a mediation model based on two parametric models. The first is the model used in Section 6.2, which regresses arousal on offers. The second model uses a probit regression where binary variable indicates whether the offer is rejected (1) or rejected (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 only occurs when offers are larger, rather than small. 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 in Imai et al. (2010).[9] With this method we must first define the

[8]All interactions completely anonymous.

[9]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 (see Section 8).

treatment contrast. In our setting this amounts to setting two treatment conditions, t', t, that correspond to two different offer levels. We report two sets of contrast, one with $t' = 0, t = 5$ and one with $t' = 5, t = 10$. 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 due to 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 hypothesis 3.

Figures 4 and 5 plot the result for the BigShift (top) and SmallShift (bottom) conditions. In each plot, the average causal mediation effect (ACME) for both treatment conditions (dark line for t and dashed line for t'), direct effect, and total effect point estimates and confidence intervals are reported. In both plots the mediation effect under t is positive and statistically significant for the BigShift condition. 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 hypothesis 1, the total effect of offers on rejection decisions is negative, large, and statistically significant.

Controlling for offer size, higher levels of skin conductance lead to a **lower** probability of rejection. In the first stage model (regression of offer on physiological arousal) the effect of the treatment is negative **AND** in the second stage of the mediation model (rejection decision on offer and physiological arousal) the effect of physiological arousal is negative. Thus the average causal mediation effect is positive (the product of two negative effects is positive). The effect of offers that is transmitted through changes in skin conductance *increases* the probability of a rejected offer. This, however, only holds for higher offers in the 0/5 contrast.

Here, we found that higher levels of skin conductance does not necessarily predict more rejection. Rather, higher levels of skin conductance actually predicted *less* rejection. It should also be noted that this finding is robust even when offer size is controlled for. 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 this deliberate process. Despite the similarities of the ultimatum game and our bargaining game, we argue that the two games are sufficiently different by design that

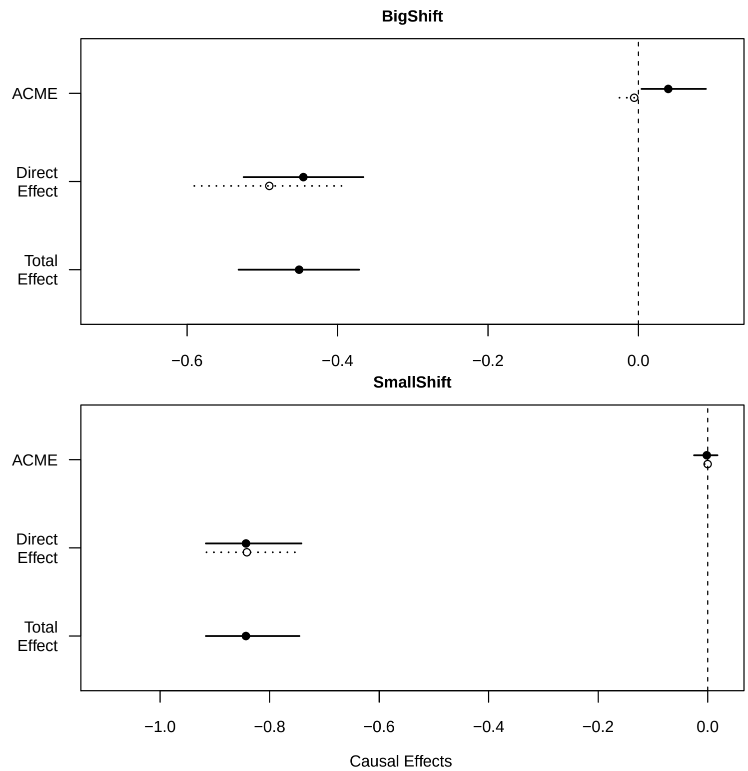

Figure 4: Mediation effects for treatment change of 0 to 5. Top row BigShift treatment and bottom row SmallShift treatment. Causal effects with 95% confidence intervals calculated through the quasi-Bayesian approximation (bootstrapping produces similar results), and allowing for clustering at the individual level.

low offers did not necessarily led to homogeneous outcomes. In particular, the negative effect of physiological arousal on rejection rates is indicative of the possibility that aversive arousal impairs one's ability to engage in strategic thinking, thus making it difficult for the responders to choose the optimal strategy of rejection.

7 Discussion & Conclusion

Explanations for the failure of bargaining and the resulting conflict have centered heavily in recent years on so-called 'commitment problems' that arise from shifting bargaining power. Because of the inherent difficulty in credibly committing oneself not to take advantage of a rival in the future, bargaining under conditions of shifting power are likely to result in violent conflict. This explanation is compelling, but makes no attempt to

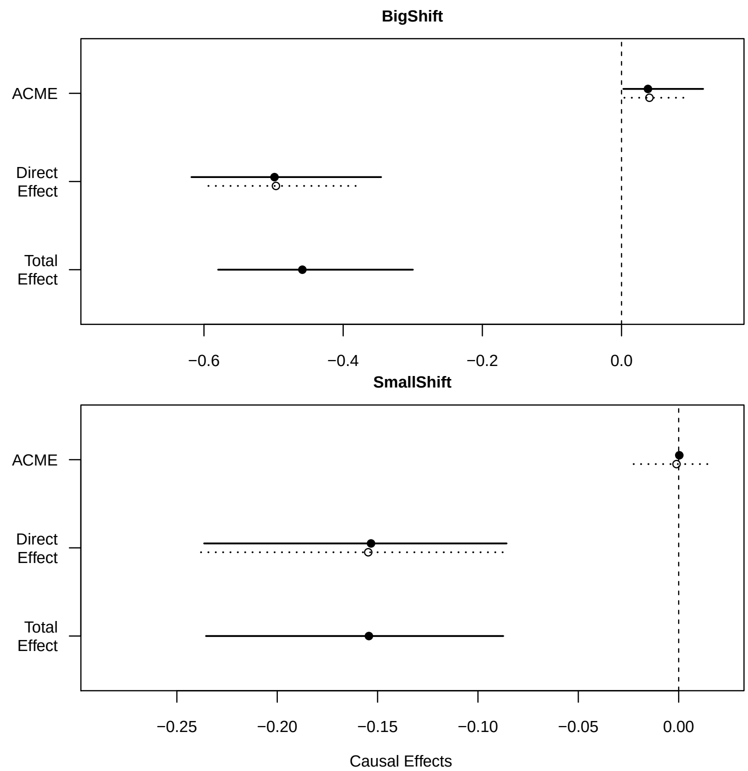

Figure 5: Mediation Effects for treatment change of 5 to 10. Top row BigShift treatment and bottom row SmallShift treatment. Causal effects with 95% confidence intervals. Causal effects with 95% confidence intervals calculated through the quasi-Bayesian approximation (boot strapping produces similar results) and allowing for clustering at the individual level.

connect to what individual, human, decision-makers might experience.

We posit a theory of 'aversive arousal' that is activated under conditions of large power shifts (but not, it should be noted, under stable conditions). While actors who face the prospect of substantially losing power should reject all offers, we show that offers can trigger physiological reactions which may impair strategic reasoning that is part of the deliberative, 'System II', process identified by psychologists. Nevertheless, we also provide indirect evidence of System II thinking, in that for many offers rejection decisions were more likely in the BigShift condition. This behavior is consistent with the predictions of the rational choice model.

We have brought together a somewhat unusual combination of formal modeling to derive hypotheses, 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. We also demonstrate how the role of physiological responses may differ depending on the type of game (e.g., previous studies look almost exclusively at the Ultimatum game). A richer understanding of the relationship between emotion and decision-making in different incentive environments would certainly be valuable (see Myers and Tingley, tted; Lee et al., aper). Finally, 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.

Of course, any successful experiment inevitably generates more questions than it answers. In this case, we focus on deliberative versus intuitive thinking. However, a large literature on emotions and bargaining suggests that specific emotions (e.g., anxiety, anger) may be an important component of the process we describe in our study. Just as importantly, one might start to add layers of complexity to the bargaining game we used in order to more closely mimic important aspects of the international system, such as the presence of international institutions that may under some conditions alleviate conflict (e.g., by preventing shifts in offers from one period to the next).

8 Appendix

Figure 6 presents results from using linear models for both the mediator and outcome models. Treatment mediator interaction in outcome model is included.

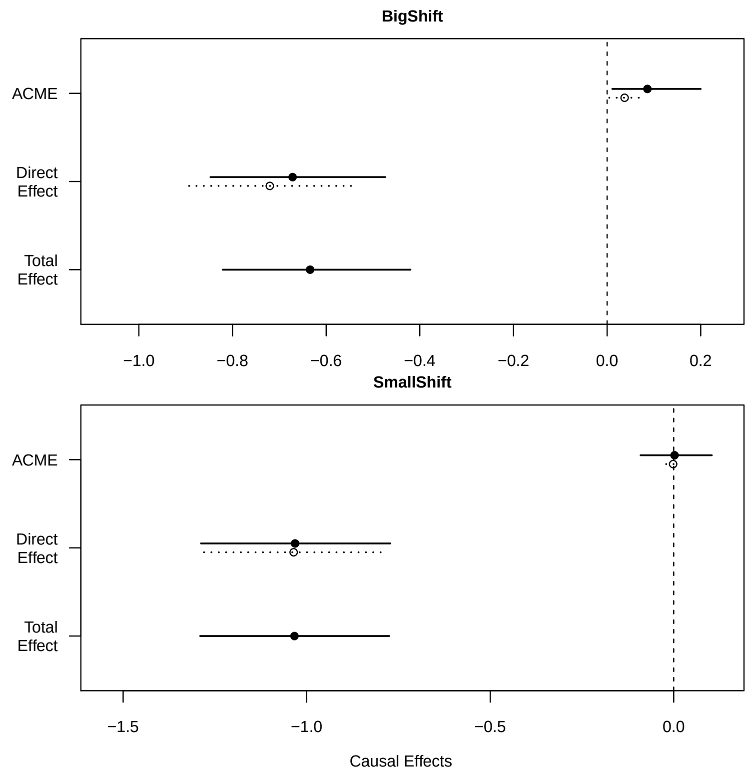

Figure 6: Mediation Effects for treatment change of 5 to 10 using linear models for both first and second stage. Top row BigShift treatment and bottom row SmallShift treatment. Causal effects with 95% confidence intervals.

References

Bechara, A., Damasio, H., Tranel, D., and Damasio, A. (1997). Deciding advantageously before knowing the advantageous strategy. *Science*, 275(5304):1293–1295.

Ben-Shakhar, G. (1985). Standardization within individuals: A simple method to neutralize individual differences in skin conductance. *Psychophysiology*, 22(3):292–299.

Blount, S. (1995). When social outcomes aren? t fair: The effect of causal attributions on preferences. *Organizational behavior and human decision processes*, 63(2):131–144.

Bush, L., Hess, U., and Wolford, G. (1993). Transformations for within-subject designs: A monte carlo investigation. *Psychological Bulletin*, 113(3):566–579.

Butler, C. K., Bellman, M. J., and Kichiyev, O. A. (2007). Assessing power in spatial bargaining: When is there advantage to being status-quo advantaged? *International Studies Quarterly*, 51(3):607–623.

Camerer, C. (2003a). *Behavioral Game Theory*. Princeton University Press, Princeton.

Camerer, C. (2003b). Behavioural studies of strategic thinking in games. *Trends in Cognitive Sciences*, 7(5):225–231.

Christie, M. (1981). Electrodermal activity in the 1980s: a review. *Journal of the Royal Society of Medicine*, 74(8):616–622.

Civai, C., Corradi-Dell Acqua, C., Gamer, M., and Rumiati, R. (2010). Are irrational reactions to unfairness truly emotionally-driven? dissociated behavioural and emotional responses in the ultimatum game task. *Cognition*, 114(1):89–95.

Dane, E. and Pratt, M. (2007). Exploring intuition and its role in managerial decision making. *Academy of Management Review*, 32(1):33–54.

Epstein, S. (1994). Integration of the cognitive and the psychodynamic unconscious. *American psychologist*, 49(8):709–724.

Epstein, S., Pacini, R., Denes-Raj, V., and Heier, H. (1996). Individual differences in intuitive–experiential and analytical–rational thinking styles. *Journal of personality and social psychology*, 71(2):390.

Evans, J. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, 59:255–278.

Evans, J. and Over, D. (1996). *Rationality and Reasoning*. Psychology Press, United Kingdom.

Fearon, J. (1995). Rationalist explanations for war. *International Organization*, 49(3):379–414.

Fearon, J. (2004). Why Do Some Civil Wars Last So Much Longer than Others? *Journal of Peace Research*, 41:275–301.

Freud, S. (1900). *The interpretation of dreams (Standard Edition, Vol. 4)*, volume 1900.

Gardner, W., Gabriel, S., and Diekman, A. B. (2000). In Cacioppo, J., Tassinary, L., and Berntson, G., editors, *Handbook of psychophysiology (2nd ed.)*, pages 643–664. New York, NY, US: Cambridge University Press.

Gigerenzer, G. and Regier, T. (1996). How do we tell an association from a rule? comment on sloman (1996). *Psychological Bulletin*, 119(1):23–26.

Gilinsky, A. and Judd, B. (1994). Working memory and bias in reasoning across the life span. *Psychology and Aging*, 9(3):356–371.

Greif, A., Milgrom, P., and Weingast, B. R. (1994). Coordination, commitment, and enforcement: The case of the merchant guild. *The Journal of Political Economy*, 102(4):pp. 745–776.

Grimm, V. and Mengel, F. (2011). Let me sleep on it: Delay reduces rejection rates in ultimatum games. *Economics Letters*.

Haidt, J. (2001). The emotional dog and its rational tail: a social intuitionist approach to moral judgment. *Psychological Review*, 108(4):814–834.

Hassin, R., Uleman, J., and Bargh, J. (2004). *The new unconscious*. Oxford University Press, New York.

Imai, K., Keele, L., and Tingley, D. (2010). A general approach to causal mediation analysis. *Psychological Methods*, 15(4):309–334.

Imai, K., Keele, L., Tingley, D., and Yamamoto, T. (2011). Unpacking the black box of causality: Learning about causal mechanisms from experimental and observational studies. *American Political Science Review*, 105(4):765–789.

James, W. (1890). *The principles of psychology*. Henry Holt and Co.

Kahneman, D. (2003). A perspective on judgment and choice: mapping bounded rationality. *American psychologist*, 58(9):697.

Lambourne, K. and Tomporowski, P. (2010). The effect of exercise-induced arousal on cognitive task performance: A meta-regression analysis. *Brain research*, 1341:12–24.

Lee, J. J., Bhanot, S., and Bruder, M. (Working Paper). How does emotion shape social preferences: Emotion as antecedents and consequences of experimental games.

Lieberman, M. (2003). Reflexive and reflective judgment processes: A social cognitive neuroscience approach. In Forgas, J., Williams, K., and von Hippel, W., editors, *Social Judgments: Implicit and Explicit Processes*, pages 44–67. Cambridge University Press, United Kingdom.

McBride, M. and Skaperdas, S. (2012). Conflict, settlement, and the shadow of the future. *working paper*.

McDermott, R., Johnson, D., Cowden, J., and S., R. (2007). Testosterone and aggression in a simulated crisis game. *The ANNALS of the American Academy of Political and Social Science*, 614:15–33.

McDermott, R., Tingley, D., Cowden, J., Frazzetto, G., and Johnson, D. (2009). Monoamine oxidase a gene (MAOA) predicts behavioral aggression following provocation. *Proceedings of the National Academy of Sciences*, 106(7):980–994.

Myers, D. and Tingley, D. (Submitted). The influence of emotion on trust.

Osman, M. (2004). An evaluation of dual-process theories of reasoning. *Psychonomic Bulletin & Review*, 11(6):988–1010.

Oxley, D. R., Smith, K. B., Alford, J. R., Hibbing, M. V., Miller, J. L., Scalora, M., Hatemi, P. K., and Hibbing, J. R. (2008). Political Attitudes Vary with Physiological Traits. *Science*, 321(5896):1667–1670.

Pillutla, M. and Murnighan, J. (1996). Unfairness, anger, and spite: Emotional rejections of ultimatum offers. *Organizational Behavior and Human Decision Processes*, 68(3):208–224.

Powell, R. (2006). War as a commitment problem. *International Organization*, 60.

Rosen, S. (2005). *War and Human Nature*. Princeton University Press, Princeton.

Sanfey, A., Rilling, J., Aronson, J., Nystrom, L., and Cohen, J. (2003). The neural basis of economic decision-making in the ultimatum game. *Science*, 300(5626):1755.

Sequeira, H. and Roy, J. (1997). Neural control of electrodermal activity. In Jordan, D., editor, *Central Nervous Control of Autonomic Function, Vol. 11*, pages 259–293. Harwood Academic Publishers.

Simmons, B. (1994). *Who Adjusts? Domestic Sources of Foreign Economic Policy During the Interwar Years*. Princeton University Press, Princeton.

Simmons, B. A. and Danner, A. M. (2010). Credible commitments and the international criminal court. *International Organization*, 64(2).

Sloman, S. (1996). The empirical case for two systems of reasoning. *Psychological bulletin*, 119(1):3.

Tingley, D. (2011). The Dark Side of the Future: An Experimental Test of Commitment Problems in Bargaining. *International Studies Quarterly*, 55(2):521–544.

Tingley, D. and Walter, B. (2011a). Can cheap talk deter? an experimental analysis. *Journal of Conflict Resolution*, 55(6):994–1018.

Tingley, D. and Walter, B. (2011b). The effect of repeated play on reputation building:an experimental approach. *International Organization*.

Tingley, D. and Wang, S. (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):243–255.

Tomz, M. (2007). Domestic audience costs in international relations: An experimental approach. *International Organization*, 61(4):821–840.

Van't Wout, M., Kahn, R., Sanfey, A., and Aleman, A. (2006). Affective state and decision-making in the ultimatum game. *Experimental Brain Research*, 169(4):564–568.

Yamagishi, T., Horita, Y., Takagishi, H., Shinada, M., Tanida, S., and Cook, K. (2009). The private rejection of unfair offers and emotional commitment. *Proceedings of the National Academy of Sciences*, 106(28):11520–11523.