
Losing Face and Sinking Costs: Experimental Evidence on the Judgment of Political and Military Leaders

Jonathan Renshon

Abstract Status has long been implicated as a critical value of states and leaders in international politics. However, decades of research on the link between status and conflict have yielded divergent findings, and little evidence of a causal relationship. I attempt to resolve this impasse by shifting the focus from status to relative status concerns in building a theory of status from the ground up, beginning with its behavioral microfoundations. I build on and extend previous work through an experimental study of status threats and the escalation of commitment, operationalized here as a new behavioral escalation task using real financial incentives and framed around a narrative of war and peace. I utilize a unique sample of high-profile political and military leaders from the Senior Executive Fellow (SEF) program at the Harvard Kennedy School, as well as a group of demographically matched control subjects, allowing me to evaluate the moderating effect of power on status concerns while also addressing typical concerns about external validity in IR experiments. I find strong evidence that the fear of losing status impedes decision making and increases the tendency to “throw good money after bad,” but that power aids decision making by buffering high-power subjects against the worst effects of status loss.

A large body of research in international relations (IR) has suggested the importance of status concerns for international conflict. Status—standing in a hierarchy—has been a critical component of both realist and evolutionary theories, as well as empirical work on status inconsistency, power transition theory, and social identity and dominance theory.¹ It has also featured in international political economy (IPE), often as one of the “less tangible payoffs” sought by states.²

I am especially grateful for the thoughtful comments and suggestions from Andrea Grove, Pete Hatemi, Josh Kertzer, Jennifer Lerner, Rose McDermott, Rich Nielsen, Christopher Oveis, Jon Pevehouse, Michelle Schwarze, Dustin Tingley, Jane Vaynman, and Jessica Weeks. I would also like to thank the editor, Michael Barnett, and three anonymous reviewers, for their helpful comments and Ryan Powers for his research assistance. The study was supported by two grants from the National Science Foundation (SES-02-39-637 and SES-08-20-441, PI: Jennifer Lerner) and was conducted as part of a long-term project on leadership decision making at the Harvard Decision Science Laboratory. This study was conducted in collaboration with Jennifer Lerner and Christopher Oveis.

1. See Markey 1999; Lopez, McDermott, and Petersen 2011; Midlarsky 1975; Volgy and Mayhall 1995; Organski and Kugler 1980; Gilpin 1983; Larson and Shevchenko 2010; and Heaven et al. 2006, respectively.

2. Elkins and Simmons 2004, 179. See also Kindleberger 1981; and Hafner-Burton and Montgomery 2006.

However, three major concerns emerge from a review of the literature on status and conflict. First, there is no empirical support for a causal link between status concerns and outcomes related to international conflict. The associations found in qualitative and large-N work cannot rule out the influence of some unobserved variable (such as a decline in trust) that affects both the likelihood of war and the measure that is often used as a proxy for status (diplomatic representation).³ Second, the social and perceptual nature of status has all but precluded direct measurement, leading to a reliance on imprecise proxies such as diplomatic representation and success in the Olympics. The experimentally induced status concerns used in this article represent a substantial improvement in this area. Finally, as Dafoe and colleagues note in a recent review of status in IR, scholars have scant information on what factors make status concerns more or less salient to political leaders.⁴ Or in other words, when do concerns over relative status overshadow the myriad other concerns that decision makers face in complex international environments? More broadly, what are the boundary conditions for theories of status in IR?

The latter point suggests that additional theoretical leverage can be gained by switching our focus from the effects of status itself (for example, the possession of high or low status) to relative concern for status, which is likely to provide variation that can be exploited in theory building and hypothesis testing. The lack of extant theory to guide us in understanding how status works in world politics, despite decades of interest in the subject, suggests that future work on the subject should take the approach of theories of reputation, audience costs, commitment problems, and realism by building from the “ground up” and providing empirical microfoundations.⁵

This article takes up that challenge by focusing on one specific type of status concern—the fear of losing status—that is commonly implicated in the literature. By experimentally inducing such fears in a laboratory, we can isolate a causal effect of status concerns on political behavior. By incorporating contextual variables—particularly power—we gain additional traction on the crucial issue of when and how concerns over status take center stage in world politics.

This study focuses on the escalation of commitment to a failing course of action, more commonly known as the “sunk costs” bias. This class of decisions—when to stop throwing good money after bad—has long been implicated in world politics as a potential cause of conflict.⁶ It has been empirically linked to the public’s appetite for war, irredentist disputes, great power military interventions, international negotiations, and the length of commitment to conflicts such as Vietnam.⁷ In contrast to

3. See, for example, Wohlforth 2009; Maoz 2010; Renshon 2014; and Volgy et al. 2014.

4. Dafoe, Renshon, and Huth 2014.

5. See Tingley and Walter 2011; Tomz 2007; Tingley, Lee, and Renshon 2014; and Kertzer and McGraw 2012, respectively.

6. See, for example, Goldgeier and Tetlock 2001; and Kahneman and Renshon 2007.

7. See Boettcher and Cobb 2009; Von Hippel 1996; Taliaferro 1998; Meerts 2005; and Staw and Ross 1989, respectively.

previous escalation research that has relied almost exclusively on hypothetical investment decisions, I introduce a newly developed experimental paradigm—the “Island Game”—to provide a behavioral measure of escalation, using real financial incentives and framed around a narrative of war and peace.

Providing behavioral microfoundations for understanding how status operates in world politics necessitates careful operationalization and research design. This is especially the case for experimental IR, which often relies on “convenience samples” of undergraduates and, increasingly, online labor marketplaces such as Amazon’s Mechanical Turk. Although questioning the use of college freshmen is no longer a novel complaint, it is still a valid one in some cases. Asking the typical sample of undergraduates to act “as if” they are political leaders can be problematic, since doing so exacerbates their own biases.⁸ Moreover, highly experienced elites may be more likely to adhere to traditional tenets of rational decision making, making studies that use undergraduates potentially misleading for the purposes of extrapolating to world politics.⁹ More generally, scholars must be wary of differences between traditional samples and leaders who are theoretically or empirically linked to the outcome of interest. In this case, it is obvious that student subjects are likely to differ from political leaders on the dimension of power (both the objective ability to exert control over others as well as the subjective, felt experience of authority), which in turn has been empirically linked to a host of judgment and decision-making outcomes. Critically, it is also often implicated in combination (or interaction) with status concerns.¹⁰

I assuage at least some of these concerns through the use of high-power political and military leaders drawn from the “Senior Executive Fellow” (SEF) program at the Harvard Kennedy School, along with a group of demographically matched control subjects from the Boston area. This provides natural variation on power within and between groups that should both address the typical concern about external validity in IR experiments and allow us to test out some of the more subtle inferences derived from the literature on status and war.

The treatment was the threat of losing status, commonly cited as one of the most important motivations for individuals as well as political leaders.¹¹ The outcome was an “escalation of commitment” task (using real financial incentives) in which subjects made a series of choices regarding whether to continue investing in a losing cause or walk away. I found that the threat of status loss did lead to higher escalation, though this main effect was not statistically significant. However, average treatment effects alone mask the highly divergent patterns of two different subgroups in the experiment. High-power subjects were buffered from the worst effects of losing status, and showed no significant decline in their decision-making abilities. In contrast,

8. Kertzer and Renshon 2014.

9. Hafner-Burton, Hughes, and Victor 2013. See also Carnevale, Inbar, and Lerner 2011.

10. See, for example, Larson and Shevchenko 2014.

11. See Frank 1985; and Tetlock 1985, on individuals, and Gilpin 1983, 31; and Horowitz, McDermott, and Stam 2005, 267, on leaders.

low-power individuals fared poorly when confronted with the threat of status loss and wasted considerably more of their money in a lost cause. These differences emerged regardless of how power was measured or conceptualized.

Context and Contribution

A careful examination of the interrelationship of status and power—and their attendant impact on outcomes relevant to international conflict—requires a reconsideration of some of the standard operating procedures of experimental IR, in particular the use of undergraduates as stand-ins for world leaders. I provide an overview of the experimental study of elites, focusing on the use of “special subjects” and more broadly on the common misunderstanding of external validity in political science. I also propose a heuristic—whether leaders and convenience samples differ on *theoretically relevant* attributes with respect to the outcome of interest—for gauging when elite subjects are necessary or even desirable in IR experiments. On a conceptual level, the leaders (and their low-power counterparts) used in this study provide a natural source of variation, and thus inferential leverage, in examining some of the more subtle implications of status-based theories of conflict.

Status Threats in International Politics

Before outlining the trajectory of status research in IR, I provide a conceptual map of status, a variable subject to a great deal of conceptual muddiness and disagreement over the years. I lean on two recent works on the subject, Dafoe’s article “Reputation and Status as Motives for War” and Larson et al.’s chapter in a new book, *Status in World Politics*, both of which, despite divergent approaches, largely agree on the major conceptual issues surrounding status and related concepts.¹²

For Larson and colleagues, status is “collective beliefs about a given state’s ranking on valued attributes,” and might be manifested in international politics as either membership in a “defined club of actors” or as “relative standing within such a club.”¹³ Dafoe and colleagues see status as “an attribute of an individual or social role that refers to position vis-à-vis a comparison group.”¹⁴ In sum, both see status as reflecting collective beliefs about where a given state stands (or ranks) with respect to comparison groups, and both agree that it can be about belonging to a given group or ranking in a hierarchy, but that in either case, positionality is critical.

12. See Dafoe, Renshon, and Huth 2014; and Larson, Paul, and Wohlforth 2014, respectively. See Wood 2013, for another perspective largely in agreement with the other two works.

13. Larson, Paul, and Wohlforth 2014, 7.

14. Dafoe, Renshon, and Huth 2014, 374.

The conceptual difference between status and prestige is perhaps less clear, as are the potential substantive effects of conflating the terms. Larson et al. define prestige as “public recognition of admired achievements or qualities” and note that it is similar to status in being relative, but does not connote ranking in a hierarchy.¹⁵ Similarly, Dafoe et al. argue that although the terms are often used interchangeably (and indeed, are often linked) and both have normative components, prestige differs in that it is generally more under the control of the actor than status is: “actors can seize, acquire and invest in their reputation and prestige. Status, on the other hand, is more often regarded as a function of the community.”¹⁶ However, status and prestige are often observationally equivalent, and the subtle conceptual differences are largely irrelevant for this work. “Status” is used in this article to refer to standing in a hierarchy, with a particular focus on which hierarchy (or “reference group”) is most relevant.¹⁷

Status has figured prominently in the “causes of war” literature in IR, most notably in the examination of status inconsistencies (divergences between status accorded an actor by the international community and what they “deserved”).¹⁸ Findings in this tradition were, however, mixed.¹⁹ Recent efforts to resurrect this research program using more precise theory and more appropriate methods have found evidence that heightened concerns about relative status are associated with the initiation of international conflict, though the exact mechanism—whether strategic or psychological—is still uncertain.²⁰

A similar focus on status can be found in hegemonic war theory, which argues that perceptions of prestige tend to lag behind a state’s actual power and capabilities.²¹ The notion of prestige is also tied closely to the “power transition” literature in which war is the result of differential rates of growth. Status is a key factor because it is one of the benefits denied to rising powers by the dominant hegemon that has “locked in” a hierarchy of prestige no longer reflective of material capabilities.²² More recently, “prestige politics” figured prominently in Kagan’s discourse on the origin of major wars, as well as Wohlforth’s theory of great power status competition.²³

Work on foreign policy analysis, social identity theory, and evolutionary approaches have come to similar conclusions regarding the importance of status concerns for

15. Larson, Paul, and Wohlforth 2014, 16.

16. Dafoe, Renshon, and Huth 2014, 376.

17. Status is distinguishable from honor, which—following O’Neill 2001—refers to either the fact of an individual possessing some positive characteristics (“personal” honor) or a group’s beliefs about an individual’s personal honor (“social” honor). Reputation, another related concept, refers to beliefs about traits or behavioral tendencies of an actor based on that actor’s past behavior. “Prestige” is used only to refer to previous work that explicitly used that term.

18. See Galtung 1964.

19. See Wallace 1973; and Midlarsky 1975.

20. See Maoz 2010; and Renshon *forthcoming*.

21. See Organski and Kugler 1980; and Gilpin 1983.

22. Lemke and Reed 1996.

23. See Kagan 1995; and Wohlforth 2009, respectively.

leaders.²⁴ Constructivist approaches often make the same point, some even using psychological theories of respect, recognition, and humiliation to serve as empirical foundations of their theories.²⁵ As Dafoe and colleagues write: “If there is one feature of reputations and status that scholars are in agreement upon, it is that leaders, policy elites, and national populations are often concerned, even obsessed, with their status and reputation.”²⁶

However, the methods employed thus far have made it difficult to assess the causal importance of status in world politics. For one, the research designs used to generate association in large data sets may be driven by omitted variables, and in case study research we often must take the “cheap talk” of leaders as gospel. The ephemeral nature of status, relying as it does on collective beliefs about others’ beliefs, has precluded direct measurement as well. This has left scholars to construct rough proxies based on diplomatic representation, the purchase of aircraft carriers and other arms, and even performance in the Summer Olympics.²⁷ Of course, no single methodological approach solves all of the problems enumerated, but experiments can provide an important complement to existing literature by directly manipulating and measuring the concept of interest, providing much greater confidence in our ability to make causal inferences about the role of status in world politics.

Critically, experimental evidence from psychology already suggests that the threat of losing status has detrimental effects on judgment and decision making. For example, uncertainty about one’s place in a status hierarchy seems to be interpreted as a direct threat and increases both the likelihood and intensity of in-group biases.²⁸ Relatedly, other researchers have found that unstable social hierarchies triggered increased activation of the amygdala, a part of the brain implicated in emotional arousal and the increased use of decision heuristics.²⁹

Power in Political Decision Making

A close reading of the literature on status and conflict suggests that the effects of status are almost always conditioned in some way on power. In status inconsistency theory, it is the discrepancy between power and status that is associated with the initiation of conflict. In power transition theory, it is a status hierarchy that no longer coincides with the balance of material capabilities that is most strongly associated with conflict. In Wohlforth’s work, it is, conversely, the balance of material capabilities that leads to uncertainty about the status hierarchy, which conflict helps to

24. See, for example, Larson and Shevchenko 2010; and Lopez, McDermott, and Petersen 2011.

25. Wolf 2011b. See also Lebow 2008.

26. Dafoe, Renshon, and Huth 2014, 381.

27. See Volgy and Mayhall 1995; Eyre and Suchman 1996; and Rhamey and Early 2013, respectively.

28. Scheepers 2009.

29. Zink et al. 2008.

resolve.³⁰ This suggests both a general agreement that power matters somehow in moderating the impact of status in world politics, but also considerable uncertainty about the actual relationship between status, power, and conflict.

Of course, actually examining the moderating impact of power requires some additional conceptual work. IR work implicates “power,” but empirically is able to measure only hypothesized determinants of power (for example, coarse proxies such as the Correlates of War’s “CINC” score).³¹ In fact, despite its centrality to theories of politics, there is little agreement about the conceptualization or measurement of power. One oft-used conceptualization relies on the notion of “power as influence.”³² In this sense, power is neither solely material nor social, but both. It is the “ability of actors to secure the desired outcomes or prevent the occurrence of undesired ones.”³³

This conceptualization suggests possible ways of operationalizing power in the lab. In dyadic, or *k*-adic interactions, the relative levels of “power” can be incorporated into the structure of the experiment, for example, by changing the odds of winning a showdown with another player.³⁴ In these cases, power is operationalized at the individual level, but we are still relying on a hypothesized connection between actual power (odds of winning) and perceived power (how players interpret those odds). This implies that we can play to the strengths of experimental methodology by focusing on the way that actual material power affects decision making: through the instantiation of low- and high-power “mindsets” that, in turn, affect judgment and behavior. The status-based theories of conflict listed in this study, which also implicate material capabilities and involve psychological mechanisms, implicitly assume that those capabilities—whether it is a balance, imbalance, or declining or rising power—are assessed and filtered through leaders’ perceptions. This dovetails with much of the more subtle realist works in IR, which emphasize subjective assessments of power and threat over the objective balance of capabilities.³⁵

So what are those consequences of power mindsets for judgment and decision making? Despite an explosion of studies on this question, social psychological research on this topic does not suggest a clear answer. In some cases, power can have salutary effects, for example by facilitating goal-directed behavior.³⁶ Some of its effects are harder to pigeon-hole as beneficial or pernicious, and are simply context-dependent. For example, the powerful tend to be less influenced by situational factors (such as situational cues) and less affected by loss aversion.³⁷

30. Wohlforth 2009.

31. Singer 1988.

32. Baldwin 1979.

33. Maoz 1989, 240.

34. As in Tingley, Lee, and Renshon 2014.

35. For example, Wohlforth 1993.

36. Magee, Galinsky, and Gruenfeld 2007.

37. See Galinsky et al. 2008; and Inesi 2010, respectively.

However, negative effects abound as well. Power can lead to overconfident decisions and economic losses, as well as impel subjects to discount advice, leading to an overall decrease in accuracy compared with their less powerful peers.³⁸ It has destructive implications for social relations because it is associated with the objectification of others and prejudice toward outgroups.³⁹ Finally, it can exacerbate other common decision biases, for instance, by increasing illusory control over events and outcomes.⁴⁰

This study centers on the escalation of commitment, but the evidence for this class of decisions is also unclear. Power increases risk taking, but this tendency is moderated by an individual's power motivation (how much they care about increasing power).⁴¹ In fact, high scores on the power motivation scale predict more conservative decisions, rather than riskier ones. These findings are buttressed by research in neurobiology showing that powerful individuals who have relatively high levels of basal testosterone (T), generally associated with the pursuit of status and power, were more conservative than their high T, low-power counterparts.⁴² In other words, there is no clear prediction for how power should affect escalation tendencies in the present study.

Power on the individual level can be operationalized in a variety of ways. In some cases, such as Guinote et al., power was manipulated directly through priming (in much the same way that status is here).⁴³ In others, such as Smith et al., power was conceptualized and measured as a stable individual difference measure (for example, the "sense of power" scale described in Anderson and Galinsky's work).⁴⁴ Some works use both approaches in combination to triangulate the effects of power.⁴⁵

We built on these approaches by constructing two measures, one objective and one subjective. This follows from the observation of Anderson et al. that (1) power is composed both of individuals' actual authority (their ability to control outcomes and other actors) and their own sense of power and (2) these dimensions of power require separate measurement.⁴⁶ In IR, measurement issues force us to focus almost exclusively on the former aspect, measured indirectly and somewhat crudely as state capacity, despite our awareness that the relationship between actors' "material capacity" and their ability to control outcomes (let alone their "sense of power") is neither monotonic nor linear.⁴⁷

38. See Fast et al. 2011; and See et al. 2011.

39. See Gruenfeld et al. 2008; and Guinote, Willis, and Martellotta 2010, respectively.

40. Fast et al. 2009.

41. See Anderson and Galinsky 2006, on risk taking, and Maner et al. 2007, on the moderating influence of power motivation.

42. Ronay and Von Hippel 2010.

43. Guinote, Willis, and Martellotta 2010.

44. Smith, Wigboldus, and Dijksterhuis 2008. On the sense of power scale, see Anderson and Galinsky 2006.

45. For example, Fast et al. 2011.

46. Anderson, John, and Keltner 2012.

47. For more on this, see Maoz 1989, 247.

To capture this distinction, and as part of a long-term project on leadership and decision making, our research team constructed two additional measures, one objective and one subjective.⁴⁸ Previewing later results, I found that both objective measures (that serve as a stand-in for “capacity”) and more subjective measures (that more closely align with individuals’ feelings of possessing power and authority) are closely correlated in this study, and that both have similar effects. This operationalization of power will be novel for most quantitative IR scholars, but focusing on the perception and feeling of power as experienced by individual leaders has a long tradition in both theoretical and empirical work.⁴⁹ Rose, for example, in one of the defining statements of neoclassical realism, argued that “foreign policy choices are made by actual political leaders and elites, and so it is their perception of relative power that matters, not simply relative quantities of physical resources or forces in being.”⁵⁰

Powerful Elites: Experimentation in IR

Because of the importance of power as a moderating factor in most status-based theories of international conflict, any experiment intended to demonstrate a causal effect of status on a conflict-relevant outcome must grapple with this complex interdependence. The experimental approach is uniquely suited to investigate how status and power affect individuals at the most micro level by focusing not on “objective” conditions in the international system—which are filtered through slow-moving and biased information processing systems—but rather on the more direct, felt experience of having (or losing) status and power. However, doing so in an experimental context requires rethinking the standard operating procedures of experimental political science, which often rely on subject pools composed of college students. I provide an argument for how to understand the debate surrounding “external validity” in the context of experimental political science.

To begin, a broad point about external validity, lab experiments, and the growing literature in experimental IR. Simply put: external validity—the extent to which a given result is generalizable to alternative contexts, populations, and measurement strategies—is not a property of individual experiments.⁵¹ Or, as McDermott puts

48. These scales are discussed in more detail in Sherman et al. 2012 and 2013.

49. On the perception of power in IR theory, see Morgenthau 1948. For an example from empirical IR, see Brooks and Wohlforth 2006 on declining power as experienced by Gorbachev during the late 1980s.

50. Rose 1998, 147.

51. Morton and Williams 2010, 255, define external validity much more narrowly as “the approximate truth of the inference or knowledge claim for observations beyond the target population studied.” It is, in other words, specific to generalizing to alternative populations, while they describe the concern with whether “methods, materials and settings of the research are similar to a given target environment” as “ecological validity.” Bundling these two concerns together, as I do in this article, is conventional in IR contexts, but also fits more closely with standard definitions, such as Campbell and Stanley 1963; and McDermott 2002b.

it: “external validity follows, as replications across time and populations seek to delineate the extent to which ... conclusions can generalize.”⁵² Yet one of the most common concerns in experimental political science is that laboratory studies (and even survey experiments) that use “convenience samples” of undergraduates cannot tell us very much about elite decision making.⁵³ As Kam aptly noted, many political scientists use a “simplistic heuristic” in which student samples prevent generalization to other contexts.⁵⁴

These concerns are neither unique to political science, nor particularly new. Psychologists have long worried that conducting the majority of the field’s research on college students would yield conclusions that lacked clear external validity. As early as 1946, McNemar observed that “the existing science of human behavior is largely the science of sophomores.”⁵⁵ Concerns over this “narrow data base” have not changed a great deal in the ensuing years.⁵⁶ In the context of IR, this dilemma is highlighted by Kertzer and Renshon, who found that asking students to “act as if” they were leaders simply caused them to act more like themselves.⁵⁷

As a productive response to these concerns, social scientists have taken to replicating findings using different populations. The results have been mixed. For example, Glaser et al. found that professional traders were just as likely as “lay people” (even more in some circumstances) to exhibit biased overconfidence.⁵⁸ In other cases, results were either moderated or wholly different. Weather forecasters, for example, who receive immediate and unambiguous feedback, do not exhibit overconfidence, and political officials and military officers seem to employ different decision strategies.⁵⁹

Political scientists have gradually begun to recruit more “elite” subjects, such as military officers, and have found significant differences between those samples and the standard student pools.⁶⁰ Fehr and List, for example, found that CEOs were both more trusting and more trustworthy than students in experimental settings.⁶¹ Alatas et al. found that Indonesian civil servants exhibited different patterns of beliefs and tolerance for corruption than Indonesian students.⁶²

Logistical problems have slowed this useful area of research, however, because scholars have found it difficult to find a suitable number of “special subjects.” Moreover, virtually no experiments have used high-level political decision makers.

52. McDermott 2011, 28.

53. Barabas and Jerit 2010.

54. Kam 2007. See also the discussion in Druckman and Kam 2011.

55. McNemar 1946.

56. See Sears 1986; and Gordon, Slade, and Schmitt 1986.

57. Kertzer and Renshon 2014.

58. Glaser, Langer, and Weber 2005.

59. See Tyszka and Zielonka 2002; and Mintz, Redd, and Vedlitz 2006, respectively.

60. For studies that employ military officers, see Mintz 2004; and Mintz, Redd, and Vedlitz 2006. For more on the differences between elite subjects and student pools, see Hafner-Burton et al. 2012; and Hafner-Burton, Hughes, and Victor 2013.

61. Fehr and List 2004.

62. Alatas et al. 2009.

Even an article titled “How Politicians Make Decisions: A Political Choice Experiment” used professors with doctorates in economics as surrogates for political leaders.⁶³ This is flattering for professors, but common sense suggests that they may not be the best stand-ins for high-level political and military leaders.

Of course, it is neither necessary nor desirable to use only political leaders as subjects, both because logistical problems make this approach extraordinarily costly, and because in many cases concerns about external validity and generalizing are either overstated or premature.⁶⁴ But even though “special subjects” are not necessary for every political science experiment, they may be particularly so when the target population differs on attributes that are theoretically relevant for a given study, such as experience, familiarity with the decision context, or age.⁶⁵ This is often the case in IR, where it is obvious that we must address the significant differences that separate undergraduates from real-world leaders.

One dimension that separates those two populations is that high-level political leaders seem likely to differ tremendously on power (the ability to exert control over others as well as the subjective beliefs associated with that authority). Despite all the many differences—age, experience, regime type—among political leaders, they are all far more powerful as a general rule than any undergraduate. This suggests that a study such as this one, focused on status and power, is one instance where a typical convenience sample of undergraduates would be particularly ill-suited both to test the relevant hypotheses as well as to generalize to high-level decision making.

A second, related concern often invoked in response to IR experiments is that even if we were able to use high-ranking political leaders as subjects, they would behave differently in the lab than they would in real-life situations. There are two reasons this might occur: either subjects would not truly engage in the task in the lab, and so we would not observe the actual decision process as it would happen in the real world, or other factors in the real world (and unaccounted for in the experiment) might induce different results.

The former rationale, the lack of “high stakes” in the lab (compared with the White House’s Situation Room), is the factor most commonly implicated as preventing us from learning from experiments. However, 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 with no financial incentives) that make a difference, not marginally increasing stakes that are already present in the experiment.⁶⁶ This is, in fact, exactly why the current experiment

63. Fatas, Neugebauer, and Tamborero 2007.

64. For a similar argument, see McDermott 2002a.

65. See Hafner-Burton, Hughes, and Victor 2013; Mintz 2004; and Horowitz, McDermott, and Stam 2005, respectively. An analogy can be made here to omitted variable bias in large-N work: scholars know that most statistical models are incomplete, but what matters is whether that incompleteness is correlated with something that is relevant for the model.

66. See Camerer and Hogarth 1999; Holt and Laury 2001; and Hertwig and Ortmann 2003.

was designed to incorporate (moderate) monetary incentives.⁶⁷ More broadly, I believe the “mundane realism” of the task to be appropriately high, and discuss this in greater detail in the description of the protocol.

On the latter question, it is always possible that there are other factors (for example, emotions, time pressure, stress, group dynamics) unaccounted for in an experiment that might make a difference in the real world. It is worth remembering that external validity is built up over time by replication and extension, and that any given experiment can only do so much without losing the very advantages that experimental methods and randomization provide.

Buttressing our confidence, there is a long tradition of experimental results generated by “trivial” tasks in laboratory settings that generalize to people in the “real world” operating in much more complex situations. This is the case with aggression, anchoring, prospect theory, and more recently in political science with negative advertising and turnout as well as findings related to immigration and Medicare.⁶⁸ These results stand as testament to Verba’s warning that “the experimental model does not need to ‘look like’ the real world. What is important is the question of whether it operates like the real world in the respects that are relevant to the study at hand.”⁶⁹

The “Levels of Analysis” Problem in Status Research

One final concern often invoked regarding experimental IR concerns the “levels-of-analysis” problem.⁷⁰ In fact, this is typically a bundle of intertwined concerns related to ecological inference (inferring individual preferences/behavior from aggregate data), assumptions about the aggregation of preferences (using data from a “lower” level of analysis to explain state- or system-level outcomes), the unit of analysis where measurement takes place and whether a given empirical test “matches” the relevant theory exactly or provides microfoundations (that is, testing out microlevel implications of higher-level theories). In some cases, theories being tested experimentally are already directed and measured at the individual level, in which case, these concerns are largely moot. However, in the cases that do not meet this standard, experimentalists must provide an account of how to think about the “micro-macro” link if their results are to hold implications for our understanding of international politics. So, at what level of analysis are status concerns implicated?

In fact, status is conceptualized and measured at varying levels of analysis. The confusion that sometimes results from this stems from the additional complexity of

67. The relationship between incentives and performance does not appear to be entirely monotonic when incentives are extremely high. See Ariely et al. 2009.

68. See Anderson and Bushman 1997; Engh and Mussweiler 2006; Camerer 2004; Ansolabehere, Iyengar, and Simon 1999; and Barabas and Jerit 2010, respectively. For an excellent discussion of this issue, see Levitt and List 2007.

69. Verba 1964, 502. See also Kelman 1965, 598.

70. Singer 1961.

status concerns, which require specification of the reference group (“who is threatening status of the actor?”) as well as whether it is group or individual status that is at stake (“whose status is threatened?”). Many extant theories of status concerns in IR either directly concern individual and/or group-level phenomena or have indirect implications for that level that are amenable to experimental tests.

For example, status inconsistency theory is almost always measured at the state or system level, while the theory underlying it (based on the “frustration-aggression” hypothesis) is individual-level.⁷¹ In a representative work, Volgy and Mayhall describe a theory in which status is measured at the system level for practical purposes, but status matters insofar as it is “salient for decision makers.”⁷² Theories of hegemonic war and power transition differ from this because they take place almost wholly at the state and system level. However, even these theories often implicate individual-level phenomena, such as perceptions of threat or credibility.⁷³

As for whose status is threatened in theories of status in IR, a fair reading of the literature suggests that it varies considerably. In Doran et al., it is the status of the nation/group that is implicated in the theory, but the empirical foundation is provided via a survey of individuals’ attitudes regarding status and threat.⁷⁴ In much other work on the role of status in IR theory, the focus is exclusively on how individuals both strive for status and react when it is threatened.⁷⁵ Social identity theory (SIT) is one particularly promising framework for examining status concerns, but contains some subtle features that belie an easy answer for whose status is really at stake. Wohlforth’s concern, for example, is “whether Great Powers care about status,” but the underlying theory (SIT) is targeted at the group level and much of the evidence relies on perceptions of key elites.⁷⁶ Larson and Shevchenko use the same group-based theory (SIT) to focus on “*states*’ concerns about their relative status,” but again they draw evidence from individual actors’ perceptions of key events and trends.⁷⁷

Although IR scholars sometimes consider these “group-level” theories, they draw on data from the individual level and base the theoretical framework on the assumption that individuals’ status goals will be transmitted through collective identity to group behavior, and that threats to the group’s status will be transmitted the opposite way, to the individual. An example of this dynamic can be found in the foundational work on the “minimal group paradigm,” in which psychologists measure the effect that being in a group (which is based on trivial or imaginary differences) has on how individuals behave.⁷⁸ In other words, the theory is “group-level” because it is the status of the group that is at stake, even if our interest is primarily in how

71. Miller 1941.

72. Volgy and Mayhall 1995, 68.

73. See, for example, Organski and Kugler 1980.

74. Doran, Hill, and Mladenka 1979.

75. See Markey 1999; and Wolf 2011a and 2011b.

76. Wohlforth 2009, 34–35.

77. Larson and Shevchenko 2010.

78. Turner, Brown, and Tajfel 1979.

individuals deal with these threats (note how different this is from how political scientists think about “group-level” theories).

These theories hinge on the core proposition (as well as decades of empirical research suggesting) that because individuals identify with groups, group phenomena are felt and can be analyzed at the individual level. In this way, the theory puts the “group in the individual.”⁷⁹ If anything, applications of these findings to politics are likely to be easier than to our individual lives, where a proliferation of overlapping identities and potential reference groups complicate the issue. Identity in international politics is in many ways much easier to conceptualize: “decision makers’ identification with the state is generally a given.”⁸⁰ In fact, a vast empirical literature on minimal groups and identity formation suggests that threats to group status are likely to be felt by individuals just as strongly as if it was a threat to their personal status position.⁸¹ This is unsurprising given the critical importance that group identity has for personal self-esteem, especially when group identity is primed or salient (such as it is likely to be for political leaders acting in their official roles).⁸²

The current project follows a similar logic to that of the studies enumerated earlier. I ask subjects to consider a decisively social event (“losing status in the eyes of your peers”), but measure its impact at the individual level. Additionally, it is the status of the individual that is threatened, a decision that was made to benefit from a tested and validated instrument. To the extent that many IR theories of status already implicate the status of individuals, this design poses no issues for interpretation. However, because some theories do implicate a combination of group and individual status, I rely on the explicit claim of Wohlforth and others who use SIT that threats to group status are, effectively, transmitted to individual leaders so perfectly that we can use psychological theories of status to understand how that individual will react to these threats.⁸³ If this is wrong—if, for example, a threat to the group’s status at magnitude x is felt by leaders as a threat to their own status at some level below x —then the results of this study would represent an overestimate of the effect, but only for situations in which individual status was not threatened, but group status was. I return to this issue in the discussion of the results, where I propose extensions that more directly address potential differences between group and individual status threats.

The current research is thus well suited to provide “microfoundations” for theories of status in world politics, many of which rely on individual and group-level mechanisms, and all of which hold at least some testable implications at the individual level. As for the relevant level of analysis for analyzing power in international relations, it is obvious that power in IR is typically a state-level characteristic, often measured as state capacity. However, as noted earlier, many influential works, recognizing that our measures of power are necessarily filtered imperfectly through the

79. Abrams and Hogg 1988.

80. Wohlforth 2009, 26.

81. Ellemers, Spears, and Doosje 2002.

82. Mercer 1995.

83. Wohlforth 2009, 37.

perceptions of individual leaders, have focused precisely on those individuals' beliefs. Thus, to the extent that material constraints affect behavior through the instantiation of beliefs about power, it follows that we should take advantage of the laboratory setting to precisely measure those beliefs. Previous experimental work has leveraged the precision inherent in experimental methods to measure power at the individual level (as individual capacity or probability of winning).⁸⁴ I provide an even finer-grained measurement of this important concept by measuring beliefs about power at the individual level.

Status, Power, and Escalation in the Lab

There are two critical features of my experimental design that strengthen its relevance for IR. The first is that the decisions have economic consequences: subjects lose actual money if they perform poorly, so there are strong incentives to pay attention and do well. Although it does not match the stakes involved in international crises, even small economic stakes can make a big difference when it comes time to generalize to situations outside the lab.⁸⁵

Second, this study uses a subject group composed of political and military leaders. This is not critically necessary in all experiments, and one might argue that even in studies relevant to politics, the typical sample of college students is a good place to start (and, of course, even relatively high-level leaders in the military and civilian bureaucracy are several steps removed from inviting the National Security Council into the lab). However, the subjects used in this study are far more representative of the political leaders we are generally interested in—and more similar on relevant dimensions—than almost any other study in experimental political science. This furnishes, at the least, some amount of increased confidence in the ability to generalize to real-world political decision making. Additionally, these high-level subjects provide natural variation on a dimension that is of theoretical relevance to the outcome and to politics more generally: power.

Procedure and Hypotheses

The seventy-seven adults in the “leadership sample” were members of the SEF program at the Harvard Kennedy School, and were overwhelmingly male (76

84. For example, Tingley, Lee, and Renshon 2014.

85. See Holt and Laury 2001; and Hertwig and Ortmann 2003. Kahneman and Tversky 1979, 265, recommend large hypothetical gambles over “contrived gambles for small stakes,” but there are obvious trade-offs in either method. I emphasized the “mundane realism” of real economic stakes to ensure that participants paid attention and tried to win. The debriefing sessions assuaged potential concerns that the money involved would be too small to induce effort; subjects were fiercely competitive, constantly checking with each other to see who won, how they did, what strategy they used, etc.

percent) and highly educated (75 percent had postgraduate educations). SEF members are typically high-ranking government employees and military personnel who possess the rank of GS14 or GS15 (for civilians) and O-5 and O-6 in military grades and generally include members of every branch of the US armed forces as well as representatives from other agencies, such as Federal Emergency Management Agency, Department of Energy, Department of Defense, and Department of State as well as several foreign governments.⁸⁶ Often, the members of the military sent to the SEF program are being screened for promotion to “flag rank” (O-7 or above). The “control population” included sixty-four adults from the Boston metropolitan area.⁸⁷ To the extent possible, they were matched to the leaders on demographic attributes, and a comparison reveals that this goal was largely accomplished.⁸⁸

The study took place at the Harvard Decision Science Laboratory in fall 2010. Subjects first completed a battery of demographic and individual difference measures. Following that, they were randomly assigned to one of three status essay conditions and then completed a decision task that measured the escalation of commitment to a failing course of action. The experimental protocol is depicted in [Figure 1](#). All subjects took part in other, unrelated studies during the sessions that generated the data used in this article.

The structure depicted in [Figure 1](#) allows us to estimate a casual effect for status concerns on escalation, but it is worth noting that the outcome itself has nothing to do with status. This structure is analogous to the way in which researchers have used the “carry-over” effects of emotions to properly isolate their causal effects: status concerns triggered by thinking about a situation that might happen in one’s own life carry over to affect what should be an unrelated decision regarding escalation in the Island Game.⁸⁹ It is with this structure, and only with this structure, that we are able to properly isolate the causal effect of status concerns on behavior.

Leadership and Power Measures. The first measure of power (LEADERSHIP GROUP) was binary and indicated whether the subject was part of the SEF course at the

86. There were not enough foreign nationals to draw meaningful inferences about cross-cultural variation, so they were excluded from the analysis.

87. Although the control group was paid a five-dollar show-up fee, leaders were not. Our research team was advised that such small compensation for their time would be insulting to SEF members, and depress participation. Once the study began, the incentives for both groups were identical, and included payment from the escalation task described later.

88. The gender distribution was similar (80 percent male, compared with 76 percent for leaders), as was the distribution of age ($mean_{leaders} = 48.5$, $mean_{controls} = 51.0$) and education ($mean_{leaders} = 4.8$, $mean_{controls} = 4.5$, scaled from 1 (high school) to 5 (postgraduate or professional degree)). They also scored similarly on the Cognitive Reflection Test ($mean_{leaders} = 0.8$, $mean_{controls} = 1.0$), a common proxy for cognitive ability scaled from 0 to 3 (Frederick, 2005) and the Wechsler Adult Intelligence Scale ($mean_{leaders} = 34.9$, $mean_{controls} = 32.3$; Wechsler 1997). Distributions are in [Figure 6](#) in [Appendix A](#) in the online appendix.

89. For carry-over effects in psychology, see Lerner, Small, and Loewenstein 2004; in political science, see Tingley, Lee, and Renshon 2014.

Harvard Kennedy School or a matched control. However, although there are reasons to believe that members of the SEF cohort are typical of the “elite leaders” of IR theory, not every subject will be a future general or cabinet secretary. In other words, some of the “control” subjects may in fact be leaders in real life (and, less likely, vice-versa). This implies that we should account for variation within—and not just between—the two different groups. To address this, we constructed two additional measures, one objective and one subjective.⁹⁰ This dichotomy follows from Anderson et al.’s observation that power is composed of two elements—individuals’ actual authority and their own sense of power—which are not always in accord with one another.⁹¹

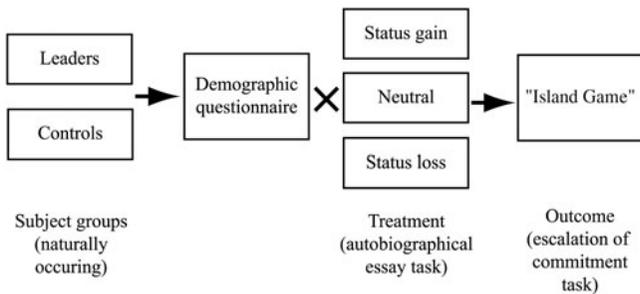


FIGURE 1. *Experimental protocol*

The first—the OBJECTIVE AUTHORITY SCALE—was a composite of questions measuring how many people were subordinate to (or managed by) the subject over the course of their career. Table 1 depicts a powerful group of political and military leaders, combined with a control group that is significantly less powerful than the leaders. The second—the SUBJECTIVE POWER POSITION SCALE—was designed to capture the subjects’ subjective evaluation of how powerful they are in their daily lives.⁹²

One concern would be a lack of within-group variation, leading us to simply re-measure our rougher, group-based measure. Figure 2 shows the distributions of both scales by membership in the LEADERSHIP GROUP. They demonstrate two important patterns. First, as one would expect, leaders generally score higher on both scales. Importantly, however, the distributions do overlap significantly, capturing the low-power members of the LEADERSHIP GROUP (some of them may in fact have jobs that resemble midlevel bureaucrats more than they do elite decision makers) as well as the high-power members of the CONTROL group.

90. See Sherman et al. 2012 and 2013.

91. Anderson, John, and Keltner 2012.

92. More information on the scales is contained in Appendix B in the online appendix.

TABLE 1. Summary statistics for objective leadership scale

	Leaders	Controls
How many people do you manage?	48.6	7.2
How many people are subordinate to you?	344.2	9.2
What is the maximum number of people that you have managed?	175.1	69.2
What is the maximum number of people that have been subordinate to you?	566.4	65.3

Note: All numbers are means calculated from raw scores, without transformations.

Treatment: Status Threat. Each subject was randomly assigned to one of three conditions in an autobiographical essay task, making this a 2 (leader/control) \times 3 (status threat/status gain/neutral), fully crossed, between-subjects design. Participants in the STATUS THREAT condition were asked to spend several minutes imagining and writing about a situation at work in which they might lose status in the eyes of their peers.⁹³ They were further instructed to focus on how they would feel when they were worried about losing the status, but had not actually lost it yet. Subjects in the STATUS GAIN condition were asked to spend several minutes imagining and writing about a time at work when they might gain status and subjects in the NEUTRAL condition were asked to think and write about a typical day at work.

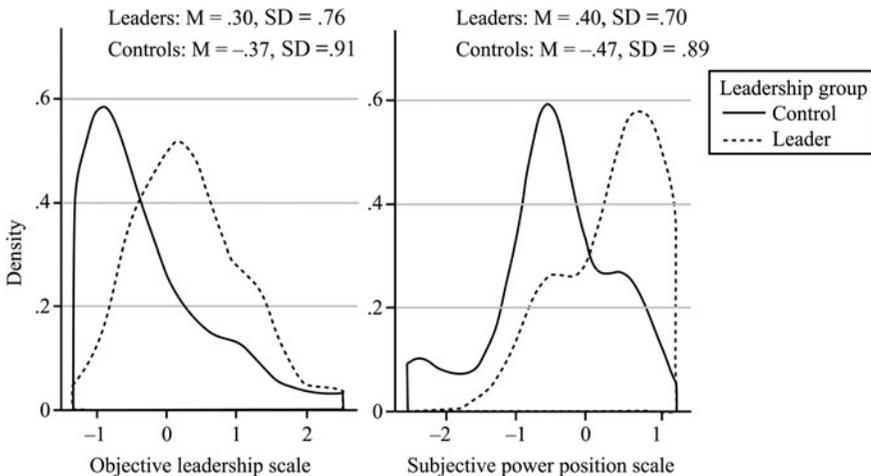


FIGURE 2. Power measures by leadership group membership

93. The status manipulation was borrowed from Pettit, Yong, and Spataro 2010. The full text of the essay instructions are contained in Appendix C in the online appendix.

Outcome Measure of Escalation: The “Island Game.” The outcome measure was the expenditure of resources in an escalation task. “Escalation of commitment” problems represent a class of decisions in which actors must decide whether to cut their losses and walk away or “double down” in an effort to recoup past losses.⁹⁴ The fixation on sunk costs has been implicated as one of the critical mistakes often made by leaders in world politics.⁹⁵

Unfortunately, the bias is as persistent as it is pernicious, proving to be surprisingly resistant to de-biasing.⁹⁶ Even replacing individuals, although intuitively appealing, does not always work, because individuals often seem to go to great lengths to justify the actions of those they replace, triggering the same escalation tendencies that the change in personnel was designed to avoid.⁹⁷ For example, even after replacing President Lyndon Johnson with a campaign premised on cutting losses in Vietnam, President Richard Nixon found it extremely difficult to actually halt the US escalation of commitment.⁹⁸ Critically, this class of decision problem has obvious applicability and relevance to domestic and international politics. Unfortunately, most experimental designs in the escalation literature use a set of tasks, such as business reinvestment decisions, that are both hypothetical and far removed from the domain of politics.

To remedy these problems, I developed a new experimental paradigm—the “Island Game”—which provides a behavioral measure of escalation using real money and framed a narrative of political decisions relating to war and peace. Participants were endowed with \$10 in the form of 100 “units” of soldiers (each valued at \$0.10) and control over 50 percent of the island, which is visually represented by a drawing of an island divided equally into two colors.⁹⁹ Each turn, they were asked if they would like to end the game, and cash in their units, or allocate a number of their units to try to capture more of the island (capturing the whole island netted a \$200 prize). If they chose the latter, they were asked how many units they would like to allocate in that particular turn. They could allocate a maximum of fifteen units or however many they had left (whichever was lower).¹⁰⁰

94. The literature on this class of problems in psychology and economics is voluminous, and focuses on sunk costs, mental accounting, emotions, prospect theory, and self-justification. A full review is outside the scope of this work, but see Staw 1981; Bazerman, Giuliano, Appelman 1984; and Wong, Yik, and Kwong 2006.

95. Walt 2010. For more examples, see Goldgeier and Tetlock 2001; Kahneman and Renshon 2007; Boettcher and Cobb 2009; Taliaferro 1998; Staw and Ross 1989. For alternative perspectives on the “irrationality” of escalation in IR, see Bowen 1987 and Friedman 2014.

96. Renshon and Renshon 2008, 527–28.

97. On “vicarious entrapment,” see Gunia, Sivanathan, and Galinsky 2009. Children are more likely to exhibit the normatively rational behavior associated with ignoring sunk costs, but this does not provide much in the way of practical advice for politicians. See Arkes and Ayton 1999.

98. Welch 2005.

99. Another, unrelated study focused on the effect of a sunk costs manipulation. However, the manipulation was most likely overshadowed by other parts of the instructions, which resulted in no main effect of the SUNK COSTS condition. Results discussed are unaffected by controlling for the sunk costs condition or interacting that condition with any of the other main variables of interest, so I collapse across the SUNK COSTS and NO SUNK COSTS conditions to gain statistical power.

100. Full instructions are contained in Appendix D in the online appendix.

The amount of the island captured each turn was determined by a weighted function such that in each turn they had a very small probability of capturing the whole of the remaining territory (and thus the grand prize) and a larger chance, weighted by the number of units they allocated, of capturing smaller pieces of territory. Each round comprised one turn by the participant and one turn by the computer (which could also capture a small bit of territory). The algorithms determining outcomes in the game are weighted such that—at the end of the computer’s turn—the subject was always in a worse position than at the beginning of the turn. Put simply, as long as subjects continued to play the game, they were pushed back into worse and worse strategic positions, steadily losing territory on the island (no subject won the “grand prize” of \$200). This is critical because we wanted to structure the game so that everyone had a similar experience: that of losing ground.¹⁰¹ The game ends when the participant “walks away” (keeping whatever money was left) or when they run out of units to use.¹⁰²

Although the scenario was made as realistic as possible, we do not need to count on the subjects feeling as though they are really making decisions about war and peace (“mundane realism”). Rather, the decision is designed to mirror, in its essence, the types of decisions that political leaders face on a daily basis: whether to continue investing resources in what appears to be a lost cause or walk away. This strategy maximizes “experimental realism” and takes advantage of our ability to isolate and identify causal factors in the lab.

Hypotheses. Although there is no extant literature on the exact question of whether the threat of status loss leads to the escalation of commitment to a failing course of action (and how power moderates this relationship), there is enough evidence on related questions to generate preliminary hypotheses, both on the direction of the effect, as well as why it is likely to occur (that is, the causal mechanism). On the first question—how status threats are likely to affect escalation—related empirical and theoretical work suggests strongly that status threats → increased escalation of commitment.

For example, Pettit and colleagues found that the prospect of losing status increased the value of status to participants.¹⁰³ If the status threat manipulation increased their value for status (in essence, how much they cared about it), and to the extent that “losing” the Island Game might be expected to lower their status

101. To avoid deception, subjects were given purposely vague instructions that they would be interacting with “another player.” The focus on escalation of commitment to failing courses of action required that subjects lost ground in every turn, unfortunately precluding the use of real dyadic interactions.

102. That subjects were not specifically prohibited from interacting following the experiment might constitute a weak accountability induction. Accountability can reduce “sunk costs” effects in some cases (Simonson and Nye 1992)—though this would likely reduce estimates of escalation across the board—but can have unpredictable effects, not all beneficial: for example, it makes a difference whether actors learn of the need to account for their actions pre- or postdecision. In the latter case, accountability can increase escalation, whereas in the former case it can attenuate it, but even then only under specific conditions. See Lerner and Tetlock 1999.

103. Pettit, Yong, and Spataro 2010.

(given that they likely anticipated discussing it with fellow participants following the study), we would expect increased expenditure of resources for those in that experimental condition. Under this framework, the increased escalation occurs because concerns about status are raised high enough that they overshadow the increased expenditure of economic resources.

Others have found that highly aversive states (such as having power without status, or focusing on financial difficulties) impair cognitive functioning.¹⁰⁴ Scheepers, for example, found that high-status groups faced with the prospect of losing status exhibited the physiological symptoms associated with threat response, which decreased functioning (in comparison with low status groups who faced the prospect of gaining status).¹⁰⁵ Under this framework, it is the “nonrational” escalation of commitment that is the focus. If status threats are aversive enough to impair cognitive functioning—by consuming precious mental resources that cannot then be utilized to combat decision-making biases—we would again expect it to lead to greater escalation of commitment. This premise relies on the folk theory that greater cognitive effort necessarily correlates with less “biased” judgment, though this is not always the case. In fact, escalation of commitment is notoriously difficult to attenuate, though it is unclear if cognitive effort and attention might accomplish that goal.

Finally, and perhaps most importantly, a large literature has demonstrated the importance that being in a “loss frame” can have on risk acceptance.¹⁰⁶ In fact, this is one of the most popular of the many explanations for the sunk costs effect.¹⁰⁷ The intuition behind this is that the shape of an actor’s value function (concave in domain of gains, convex in the domain of losses) implies that, once losses are sustained, additional losses do not add nearly enough additional disutility to outweigh the value from a large gain, unlikely as it may be. This is exacerbated by the certainty effect, which magnifies the psychological salience of certain losses (relative to what they “should be” if our decision weights correlated more perfectly with actual probabilities), increasing the desire to avoid defeat and the tendency to take greater risks to do so. Thus, we would expect subjects exposed to the prospect of losing status to escalate more in the main outcome task relative to those in either the neutral or gain conditions.

Although these theoretical and conceptual works all suggest that the threat of status loss should lead to greater escalation, they present three very different mechanisms linking the status threat to increased escalation: increased value for status, impaired cognitive functioning, and prospect theory/increased risk-taking.

The expectations regarding the moderating effect of power are not quite as clear. Anderson and Galinsky, for example, find that power seems to induce optimism and risk seeking in most individuals.¹⁰⁸ Yet, here, power did not operate in a

104. See Fast, Halevy, and Galinsky 2011; and Mani et al. 2013, respectively.

105. Scheepers 2009.

106. Kahneman and Tversky 1984.

107. For the foundational works on this, see Thaler 1980; and Arkes and Blumer 1985.

108. Anderson and Galinsky 2006.

vacuum, but rather in conjunction with the prospect of gaining or losing status. Because of this, the work of Inesi seems more relevant.¹⁰⁹ She found that powerful individuals were less sensitive to losses than low-power counterparts. This suggests that power may provide a “buffer” for powerful individuals, who are less likely to be affected by the experimental status manipulation. This is corroborated by other work, which has found converging evidence that power can protect individuals from the “press of the situation.” Powerful individuals are, for example, less likely to allow outside influences to change their stated belief, whereas less-powerful subjects exhibited impaired “executive functioning” in memory, inhibition, and planning tasks.¹¹⁰

Power Buffers Leaders Against Status Threats

How Did the Treatment Affect Feelings Related to Status?

After the initial manipulation, subjects were asked: “While completing the writing task about work, please indicate how you felt on the following scale” from 1 (low) to 7 (high) for *status*, *respect*, *powerful*, and *influential*. Figure 3A demonstrates that the STATUS THREAT essay made subjects feel as if they were significantly lower status than the other conditions. For the composite manipulation check question (which combined results for status, respect, powerful, and influential), the STATUS THREAT treatment had a significant effect relative to the NEUTRAL ($t(82.2) = -2.52$, $p < .05$) and STATUS GAIN ($t(85.4) = -4.37$, $p < .0001$) conditions. However, the STATUS GAIN treatment did not differ significantly from the NEUTRAL condition ($t(80.1) = 1.43$, $p = .16$).

Next, subjects were asked how much they valued having status. Figure 3B shows the STATUS THREAT condition to be significantly different from the NEUTRAL condition ($t(83) = -2.52$, $p < .001$). In sum, the results depicted in Figure 3 replicate the findings of Pettit and colleagues regarding the effects of the essay manipulation on feelings of status and “value for status.”¹¹¹

Did the Threat of Status Loss Increase Escalation?

Although there were strong and obvious effects of the status threat manipulation on subjects’ beliefs and feelings about status, and although those in the STATUS THREAT condition did escalate, on average, more than their counterparts in the NEUTRAL or STATUS GAIN conditions, these differences were not statistically significant.¹¹² Thus, I do not find support for the notion that status threats—on their own—led to differences in escalation patterns. However, this seems to have resulted from the strong

109. Inesi 2010.

110. See Galinsky et al. 2008; and Smith et al. 2008, respectively.

111. Pettit, Yong, and Spataro 2010.

112. Escalation by condition can be seen in Figure 8 in Appendix E in the online appendix.

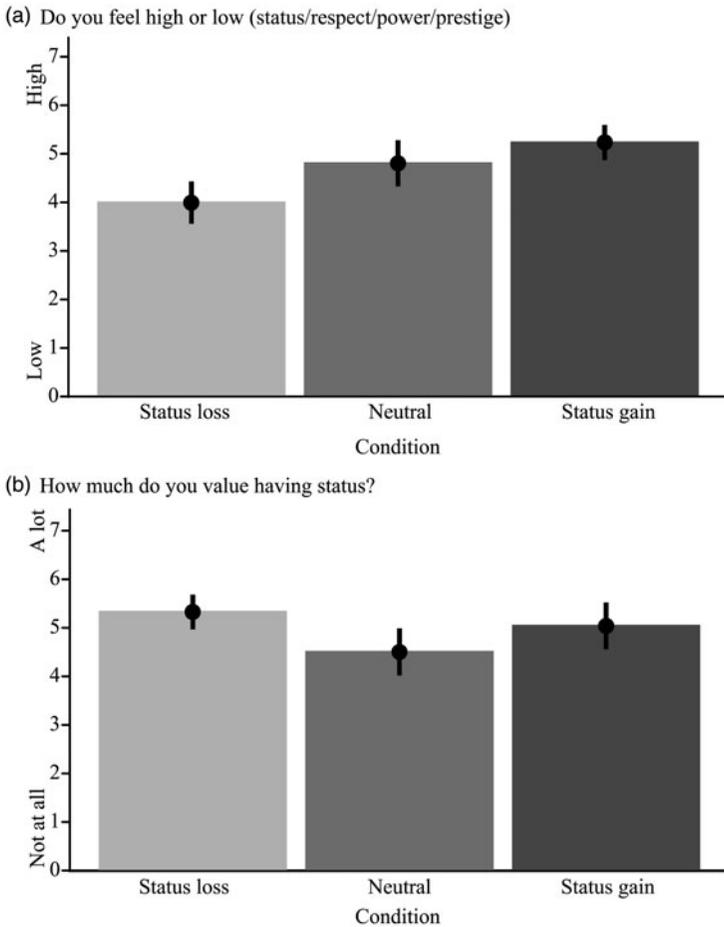


FIGURE 3. Manipulation check: Effect of essay task on status beliefs

moderating effects of power. When high- and low-power individuals are grouped together, the effects—which go in opposite directions—are washed out.

The Moderating Influence of Leadership and Power

We are left with an apparent puzzle: the status threat manipulation exerted strong effects on subjects' beliefs and feelings, but there was no main effect on escalation in the whole sample. Next, I analyze escalation in more detail by investigating the moderating effect of power. Because the nature of this work deals with political decision making, and even more specifically with high-level decisions related to war and peace, one natural line of inquiry is how our high-power subjects (the

“leadership” sample) fared in the task compared with control subjects. Or more generally, we might ask: Does power help or hurt decision making?

I estimated several regression models to shed light on how leadership affected the tendency to escalate in the primary decision task. These models include the primary manipulation (STATUS THREAT), and a measure of power plus the interaction term. In addition to the three observed measures of power—LEADERSHIP GROUP, OBJECTIVE AUTHORITY, or SUBJECTIVE POWER POSITION—I modeled power as a latent variable using polychoric factor analysis (only one factor was produced).¹¹³ All visualizations are based on linear regressions.¹¹⁴ Though the dependent variable is not normally distributed, alternative specifications using Generalized Linear Models yielded identical results.¹¹⁵

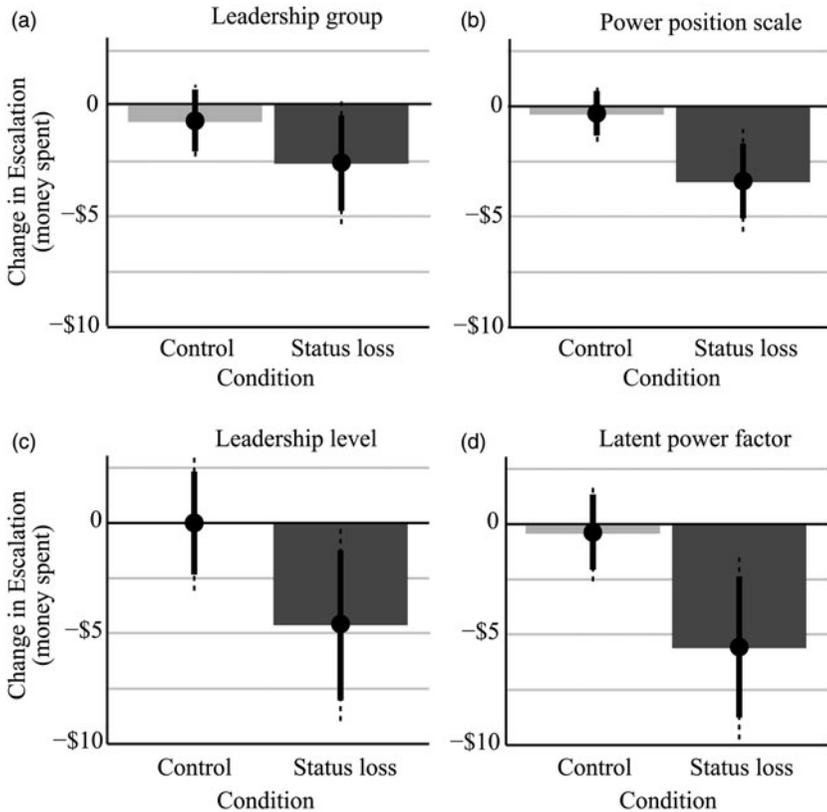
Whether power is conceptualized as membership in the “Leader” sample, as authority at one’s job (whether measured objectively or subjectively) or even as an underlying factor, the interaction term is significant (moderately so for the rougher, group-based measure, highly so for the latter three more fine-grained measures of power). This suggests strongly that the effect of status concerns on escalation depended on individuals’ level of power. However, substantive effects can be difficult to determine in regression tables, especially when considering interaction terms and conditional main effects. To more clearly illustrate this pattern, the moderating effect of leadership is plotted in Figure 4.

Figure 4 plots first differences: differences between the expected value (EV) of an outcome measure (escalation) at two different levels of an independent variable

113. All models control for the number of rounds played to rule out the alternative explanation that subjects simply played longer because they enjoyed playing the game. The advantage of this method is in keeping all observations but statistically modeling the length of playing time, while the potential concern is that it may result in a form of posttreatment bias. However, these worries are somewhat assuaged by regressing the treatment on the number of rounds played. The results were n.s. under all specifications, ruling out the possibility that part of the treatment’s effect “flowed through” the number of rounds played. The distribution of rounds played was highly skewed, with > 95 percent of the observations between zero and twenty-one and several outliers who played as many as sixty-three rounds. Appendix F in the online appendix presents results in which those few outliers are dropped and the number of rounds played is no longer controlled for statistically. The substantive interpretation of the results does not change, but this modification slightly affects levels of statistical significance for some covariates, particularly for the coarser measure of power (LEADERSHIP GROUP). Future iterations of the escalation task will constrain the number of rounds played to rule out this potential confounder.

114. The linear regressions, provided in Table 2 in Appendix F in the online appendix, also include controls for state-trait anxiety, positive and negative affect (“PANAS” inventory), gender, age, education, and whether or not the participants served in the military. On anxiety scale, see Marteau and Bekker 1992; on PANAS, see Watson, Clark, and Tellegen 1988. There is some ongoing debate about the extent to which including pretreatment controls in regression analyses of experimental data biases the results. Freedman (2008), for example, suggests that bias is likely when $N < 500$, while Green (2009) uses simulations and replications to demonstrate that the practical effect of such bias is “negligible” for $N > 20$. This makes little practical difference for the results in this case, though models with no controls are contained in Appendix G in the online appendix. There were too few women in the sample (twelve in the Control sample and eighteen in the Leader sample) to draw strong conclusions about any gender effects, but gender was controlled for in all models. Results remain the same in terms of direction and statistical significance if confined to men only.

115. See Table 2 in Appendix F in the online appendix.



Notes: Plotted values represent—for each condition—the change in expected value of outcome measure (money spent) switching from (A) control to leader; (B) mean to maximum values on SUBJECTIVE POWER POSITION scale; (C) OBJECTIVE AUTHORITY scale; and (D) LATENT POWER FACTOR. Wider, black lines are 95 percent confidence intervals while dashed lines are 99 percent confidence intervals.

FIGURE 4. *The moderating effects of power*

(power).¹¹⁶ In this case, Figure 4A shows the expected change in escalation if we switch from *Leader* = 0 to *Leader* = 1, or more plainly, from a subject recruited from the general Cambridge/Boston population, to an SEF leader. Figure 4B, C, and D show the expected change in escalation if we switch from the mean value on the SUBJECTIVE POWER POSITION, OBJECTIVE AUTHORITY, or LATENT POWER FACTOR scales to a high value.

This allows us to ask how leadership moderated the effect of status concerns on escalation. In fact, power—no matter how it was measured—greatly aids decision

116. Generated with Clarify. Tomz, Wittenberg, and King 2003.

making by buffering subjects against the pernicious effects of status threats. These plots were produced by modeling the difference between outcomes—at two different levels of the independent variable—in the CONTROL condition and then comparing it with the difference between the expected outcomes under the STATUS THREAT condition. This is thus a relatively easy way of visualizing differences-in-differences. By comparing the first differences for the CONTROL and the STATUS THREAT conditions, we get a visual representation of how power and leadership moderate the worst effects of the status threat treatment.

As Figure 4 shows, power—no matter how one conceptualizes or measures it—has no detectable effect in the CONTROL condition. In other words, power does not seem to produce the overconfidence found by some other recent works, nor does it improve decision making when acting “on its own” (there are slight differences in means, but none are statistically significant).

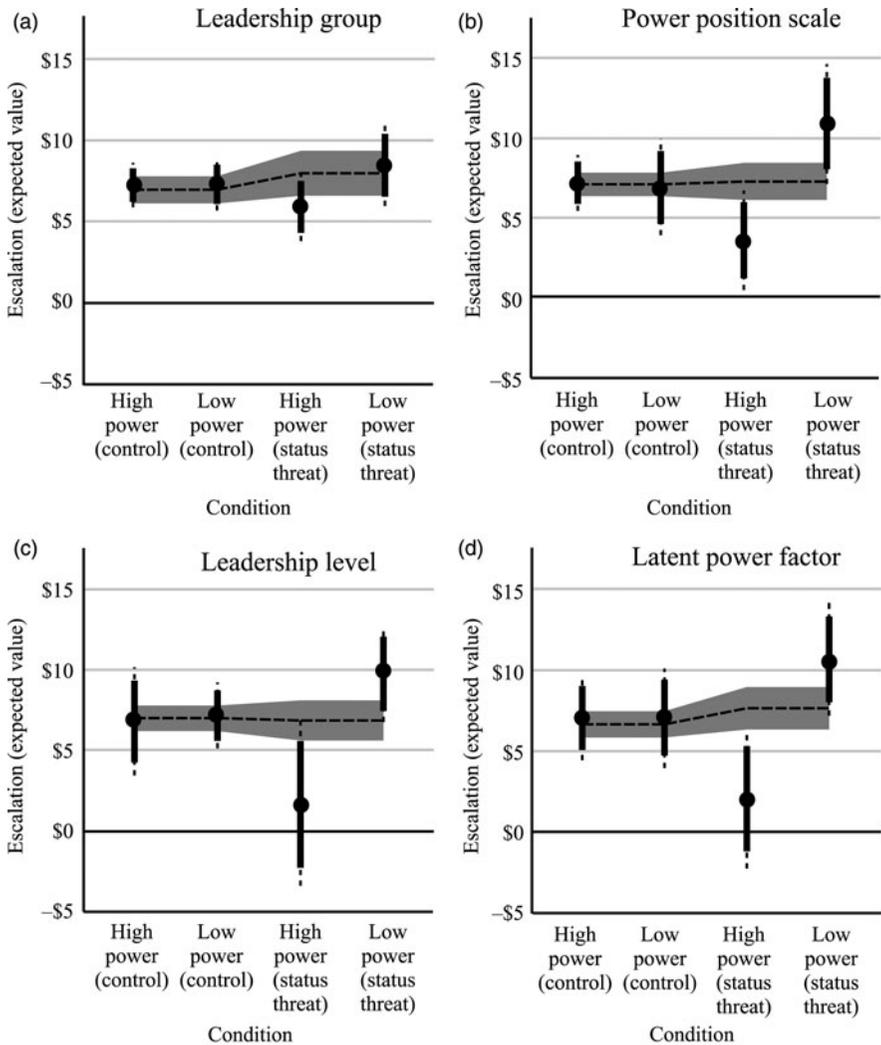
However, significant differences emerge in the STATUS THREAT condition. Where power and leadership had no discernible effect under the control condition, once status is threatened, power is strongly associated with better outcomes, in particular less escalation of commitment to what is clearly a failing policy. Powerful subjects, whether measured by LEADERSHIP GROUP (subjective or objective), authority at work, or as a latent factor, invested significantly less money in the “Island Game.” Or in other words, once status is threatened, power seems to buffer individuals against the deleterious effects of the social threat.

Although Figure 4 is helpful in depicting the moderating effects of power, it cannot disentangle two competing explanations. Either

1. Low power is increasing the effect of the status threat, or
2. The status threat is triggering some psychological mechanism for high-power individuals that leads to less escalation than expected.

In other words, did high power lead to below-average amounts of escalation (and better outcomes than even those not faced with the status threat), or did it merely prevent the “extra” amount of escalation that would be caused by the status threat if individuals were low in power? Figure 5 subdivides groups (within each experimental condition) based on levels of power to look more closely at its moderating effects. We are also controlling for as many individual-level characteristics as we can to preempt the concern that one of those characteristics co-varies with high or low power.

For each condition, estimates are presented for high (*maximum*) and low (*minimum*) levels of power. The dotted horizontal line, along with the 95 percent confidence intervals (CIs) band, represents the average amount of escalation for the CONTROL and STATUS THREAT conditions, respectively. Each point estimate (with wider, black vertical lines representing 95 percent CIs and thinner, gray lines indicating 99 percent CIs) represents the change, within the appropriate treatment condition, of switching from average amounts of power to high or low power.



Notes: Plotted values represent—for each condition—the change in expected amount of escalation at high (maximum) and low (minimum) levels of power. (A), (B), (C), and (D) represent different ways of measuring power. Dotted lines and shaded area represent the averaged amount of escalation expected for each condition plus 95 percent confidence intervals. For point estimates, wider, black lines are 95 percent confidence intervals while dashed lines are 99 percent confidence intervals.

FIGURE 5. Escalation by status condition and power

As is evident, in all four plots, escalation—for both high- and low-power individuals—in the CONTROL condition is virtually indistinguishable from average. It is in the estimates for the STATUS THREAT condition that new patterns become visible. Focusing on Figure 5A we can see that low-power subjects (here that

would mean subjects from the general population) in the STATUS THREAT condition escalate a bit more than average, but not significantly so. Similarly, high-power (the political and military leaders) subjects in the STATUS THREAT condition escalate a bit less than average, but the estimates still overlap a bit using 95 percent CIs. This fits with earlier results and common sense, all of which suggest that there is likely to be a fair amount of variation even within the LEADERSHIP GROUP such that group membership might be a noisy way of measuring power.

For more nuanced analyses, we turn to the plots in [Figure 5B, C, and D](#), which use the power scales described earlier. The patterns for the CONTROL condition in this figure are the same: just about average, irrespective of power “condition.” However, a new pattern emerges when looking at the STATUS THREAT conditions. Here we can see some (tentative) evidence that the moderating effect of power is not working by inoculating subjects against the worst effects of status threats, or by increasing the effect of the status threat for low-power individuals, but by doing both simultaneously.¹¹⁷ Put simply, within the STATUS THREAT condition, low power increases the tendency to escalate commitment whereas high power decreases that same tendency. High power, under these circumstances, emerges as a key element of high-quality decision making.

Implications for Theories of IR

One reading of this study might be that “leaders” (because they are “high-power”) are not vulnerable to the threat of status loss and its attendant effects on judgment and decision making. This would contradict previous findings, and, I believe, be a misinterpretation. Any confusion can be cleared up by referring back to the conclusions of the experiment: (1) we can identify a real causal relationship in which status concerns, randomly assigned, led to a greater escalation of commitment to a failing course of action, and (2) that effect is attenuated by power, which serves to buffer or insulate subjects.

Finding (1) is easily reconciled with existing IR literature on status and might in fact serve as an empirical microfoundation for future work. There are two complementary implications of the second finding. The first centers on power. Because power is relative, not all leaders in all situations will be in a high-power mindset that would buffer them against status threats. In fact, a brief glance at the IR literature suggests many occasions in which leaders have become convinced—sometimes correctly, other times not—that they led a state in decline.¹¹⁸ This article’s finding simply suggests that actors in a low-power mindset might be particularly vulnerable to status threats.

117. Overinterpretation of barely overlapping CIs is not advisable because they do not necessarily imply that the estimates are not significantly different.

118. See, for example, Levy 1987; and McKeown 1991.

The second implication is that we often observe situations in which states and leaders have heterogeneous responses to status threats (whether they are the kind envisioned by status inconsistency theory, power transition theory, or something else). More broadly, this fits in with the advice of Hatemi and McDermott to take heed that “not everyone reacts to the same stimulus the same way.”¹¹⁹ Put simply, there are many cases where conflict initiation (as an example) is predicted, but not observed. This article suggests that leaders’ mindsets might provide an important clue for why we see such wide variation in responses to similar situations. Why, for example, was Kaiser Wilhelm so focused on status threats during the *Weltpolitik* era when Germany’s objective situation, both in terms of status and material capabilities, hardly warranted such concern?¹²⁰ Power concerns might provide one clue to this and other historical puzzles.

Of course, simply referring to “states and leaders” elides some of the trickier questions regarding levels of analysis. I manipulated status concerns at the individual level to test their effects, inherently social and peer-referent in nature, on individual behavior. For much of the work on status in IR whose center of gravity lies between the individual and the group, this result provides an important empirical foundation. For the first time, we can cleanly identify a causal relationship between status concerns, experimentally manipulated, and outcomes of interest to scholars of international conflict.

Of course, one key question, posed by Dafoe and colleagues is: “status in the eyes of whom?”¹²¹ Or, in other words, who is the reference group? Earlier works noted that a state’s position in a hierarchy was important, but stopped short of specifying which hierarchy we should examine. Instead, states were all placed in a *de facto* global hierarchy.¹²² Whether one believes the relevant actors to be states or leaders, it seems clear that reference groups are a good deal more targeted and purposeful than previously suggested. Or, as Frank put it: status is local.¹²³ Experimental set-ups such as this one allow us to elicit beliefs about status within the relevant peer group, but building on this will require more explicit theories about how actors define their reference groups. This is not a trivial problem, and it has in no way been “solved” by the psychological literatures that IR scholars have drawn from. However, I have suggested some preliminary intuitions elsewhere, focusing on geographic proximity and major power status as two potential candidates for reference group selection.¹²⁴

A related question spurred by the “levels of analysis” problem is whether the relevant quantity is the status of the state/group or the individual. SIT suggests that this is unlikely to matter, since leaders’ identity will derive in part from their membership

119. Hatemi and McDermott 2012, 21.

120. Renshon [forthcoming](#).

121. Dafoe, Renshon, and Huth 2014, 375.

122. For example, Singer and Small 1966.

123. Frank 1985.

124. Renshon [forthcoming](#).

in larger groups (that is, the nation-state). Thus, the status concerns of leaders acting in their role as leader will in fact be the status concerns of the nation as a whole, and the peer group will be other nations, similar on relevant or salient dimensions. This may be true much of the time, but assuming this is always the case misses the opportunity to leverage potentially useful variation. There are likely to be times when the fit between a leader's individual status and their nation's status face different incentives, or become disconnected. Personalist autocracies might be one example of a regime type in which the disconnect is almost nonexistent compared with other regimes, where there are other competing aspects of identity (for example, political party). Future research might build on this by examining the different levels at which status concerns operate for leaders.

Conclusion

Although many research programs in IR invoke status as a key factor, they have all been subject to significant limitations inherent in the methods used and thus unable to provide evidence of a causal association or even precise measurement of status itself. More generally, our intuition that status matters somehow has not been matched by the construction of explicit theories of status or empirical work identifying the precise causal mechanisms through which status concerns operate.

I argued that this impasse could be resolved through a focus on the “microfoundations” of status, which I see as combining precise measurement of status itself alongside the identification and testing of specific hypotheses and causal relationships. A related innovation was the change in emphasis from the effects of status itself, to a focus on relative concern for status, and on the factors that make this concern more or less prominent for leaders.

I reported results of an experiment in which status concerns were randomly assigned prior to an “escalation of commitment” task, with real monetary stakes, in which subjects made a series of decisions about whether to continue escalating their commitment to “winning” the territory, or cut their losses and walk away. I used a unique sample of political and military leaders (as well as a group of demographically matched control subjects) that are as close as we are likely to get to the decision makers envisioned by IR theories on the critical dimension of power. This provided valuable inferential leverage in investigating the role of power in moderating status threats. Moreover, they provide an important empirical benefit: using only “control” subjects would have led to an overestimate of the effect of status concerns. I found strong evidence that power—no matter how it is measured—provides a strong buffer against the otherwise deleterious effects of status threats. In particular, high power seemed to inoculate subjects in the STATUS THREAT condition against the impaired decision making brought about by the concern over relative status.

How do these results accord with extant theories of status in IR? In some cases, for example, status inconsistency, these results provide both an empirical microfoundation for the theory and possible answers to existing empirical puzzles. For example, if

status threats causally affect behavior at the individual level, then some amount of variation in the state-level data may be explained by the degree to which leaders are differentially constrained by domestic political institutions.¹²⁵ They also suggest an answer to the question posed by Pu and Schweller: Why do some rising Great Powers (such as China) seem to consciously refrain from maximizing their status?¹²⁶ If power is a moderating factor, then this is no longer puzzling: leaders in a high-power mindset (such as those focused on a promising upward trajectory in the international system) will be less susceptible to status concerns than theories that predict status maximization all the time would suggest. Pushed further, this idea suggests some of the boundary conditions for theories of status in international relations more generally.

The limitations of the study also suggest avenues for future research. The nature of the sample did not allow for the analysis of gender, despite the likely connection between gender and susceptibility to status concerns.¹²⁷ More broadly, experimental methods and randomly assigned status concerns give us confidence in our ability to identify a causal effect (*status loss* → *escalation*) but do not necessarily tell us anything about mechanisms (*status loss* → ? → *escalation*). I noted earlier that three potential explanations exist for why status loss led to greater escalation: by increasing actors' value for status, by impairing cognitive functioning, or, à la prospect theory, by increasing risk-taking propensities. Other possibilities, such as emotions, might also play a role. For example, high power might buffer subjects through the ability of high-power subjects to regulate emotional reactions. Finally, additional research in case or large-N data that takes into account factors commonly labeled "real-world conditions"—for example, higher stakes, organizational incentives, and group dynamics—will help to provide convergent validity for the results described in this study. (Though perfectly compatible with experimental methods generally, this particular experiment did not examine all of these factors.) This future work will help to fill in critical details and deepen our understanding of the way that status and power operate in political decision making.

Supplementary Material

Supplementary material for this article is available at <http://thedata.harvard.edu/dvn/dv/jrenshon>.

125. Siverson 1995.

126. Pu and Schweller 2014.

127. Rudman et al. 2011.

References

- Abrams, Dominic, and Michael A. Hogg. 1988. *A Social Psychology of Intergroup Relations and Group Processes*. New York: Routledge.
- Alatas, Vivi, Lisa Cameron, Ananish Chaudhuri, Nisvan Erkal, and Lata Gangadharan. 2009. Subject Pool Effects in a Corruption Experiment. *Experimental Economics* 12 (1):113–32.
- Anderson, Cameron, and Adam D. Galinsky. 2006. Power, Optimism, and Risk-taking. *European Journal of Social Psychology* 36 (4):511–36.
- Anderson, Cameron, Oliver P. John, and Dacher Keltner. 2012. The Personal Sense of Power. *Journal of Personality* 80 (2):313–44.
- Anderson, Craig A., and Brad J. Bushman. 1997. External Validity of Trivial Experiments. *Review of General Psychology*, 1 (1):19–41.
- Ansolabehere, Stephen D., Shanto Iyengar, and Adam. Simon. 1999. Replicating Experiments Using Aggregate and Survey Data. *American Political Science Review* 93 (4):901–9.
- Ariely, Dan, Uri Gneezy, George Loewenstein, and Nina Mazar. 2009. Large Stakes and Big Mistakes. *Review of Economic Studies* 76 (2):451–69.
- Arkes, Hal R., and Peter Ayton. 1999. The Sunk Cost and Concorde Effects. *Psychological Bulletin* 125: 591–600.
- Arkes, Hal R., and Catherine Blumer. 1985. The Psychology of Sunk Costs. *Organizational Behavior and Human Decision Processes* 35 (1):124–40.
- Baldwin, David A. 1979. Power Analysis and World Politics. *World Politics* 31 (2):161–94.
- Barabas, Jason, and Jennifer Jerit. 2010. Are Survey Experiments Externally Valid? *American Political Science Review* 104 (2):226–42.
- Bazerman, Max H., Toni Giuliano, and Alan Appelman. 1984. Escalation of Commitment in Individual and Group Decision Making. *Organizational Behavior and Human Performance* 33 (2):141–52.
- Boettcher, William A., and Michael D. Cobb. 2009. Don't Let Them Die in Vain. *Journal of Conflict Resolution* 53 (5):677–97.
- Bowen, Michael G. 1987. The Escalation Phenomenon Reconsidered. *Academy of Management Review* 12 (1):52–66.
- Brooks, Stephen G., and William C. Wohlforth. 2006. Power, Globalization, and the End of the Cold War. In *World War I and the End of the Cold War*, edited by Gary Goertz and Jack S. Levy, 163–99. Cambridge, MA: MIT Press.
- Camerer, Colin F. 2004. Prospect Theory in the Wild. In *Advances in Behavioral Economics*, edited by Colin F. Camerer, George Loewenstein, and Matthew Rabin, 148–61. Princeton, NJ: Princeton University Press.
- Camerer, Colin F., and Robin M. Hogarth. 1999. The Effects of Financial Incentives in Experiments. *Journal of Risk and Uncertainty* 19 (1):7–42.
- Campbell, Donald T., and Julian C. Stanley. 1963. *Experimental and Quasi-experimental Designs for Research*. Boston: Houghton Mifflin.
- Carnevale, Jessica J., Yoel Inbar, and Jennifer S. Lerner. 2011. Individual Differences in Need for Cognition and Decision-making Competence Among Leaders. *Personality and Individual Differences* 51 (3):274–78.
- Dafoe, Allan, Jonathan Renshon, and Paul Huth. 2014. Reputation and Status as Motives for War. *Annual Review of Political Science* 17:371–93.
- Doran, Charles F., Kim Quaile Hill, and Kenneth Mladenka. 1979. Threat, Status Disequilibrium, and National Power. *British Journal of International Studies* 5 (1):37–58.
- Druckman, James N., and Cindy D. Kam. 2011. Students as Experimental Participants. In *Handbook of Experimental Political Science*, edited by James N. Druckman Donald P. Green, James H. Kuklinski, and Arthur Lupia, 41–57. New York: Cambridge University Press.
- Elkins, Zachary, and Beth Simmons. 2004. The Globalization of Liberalization. *American Political Science Review* 98 (1):171–90.

- Ellemers, Naomi, Russell Spears, and Bertjan Doosje. 2002. Self and Social Identity. *Annual Review of Psychology* 53 (1):161–86.
- Enough, Birte, and Thomas Mussweiler. 2006. Anchoring Effects in the Courtroom. *Journal of Applied Social Psychology* 31 (7):1535–51.
- Eyre, Dana P., and Mark C. Suchman. 1996. Status, Norms, and the Proliferation of Conventional Weapons. In *The Culture of National Security*, edited by Peter Katzenstein, 79–113. New York: Columbia University Press.
- Fast, Nathanael J., Deborah H. Gruenfeld, Niro Sivanathan, and Adam D. Galinsky. 2009. Illusory Control. *Psychological Science* 20 (4):502–8.
- Fast, Nathanael J., Nir Halevy, and Adam D. Galinsky. 2011. The Destructive Nature of Power Without Status. *Journal of Experimental Social Psychology* 48 (1):391–94.
- Fast, Nathanael J., Niro Sivanathan, Nicole D. Mayer, and Adam D. Galinsky. 2011. Power and Overconfident Decision-making. *Organizational Behavior and Human Decision Processes* 117 (2): 249–60.
- Fatas, Enrique, Tibor Neugebauer, and Pilar Tamborero. 2007. How Politicians Make Decisions: A Political Choice Experiment. *Journal of Economics* 92 (2):167–96.
- Fehr, Ernst, and John A. List. 2004. The Hidden Costs and Returns of Incentives. *Journal of the European Economic Association* 2 (5):743–71.
- Frank, Robert H. 1985. *Choosing the Right Pond*. New York: Oxford University Press.
- Frederick, Shane. 2005. Cognitive Reflection and Decision Making. *Journal of Economic Perspectives* 19 (4):25–42.
- Freedman, David A. 2008. On Regression Adjustments to Experimental Data. *Advances in Applied Mathematics* 40 (2):180–93.
- Friedman, Jeffrey A. 2014. How Cumulative Dynamics Affect Military Decision Making. Working Paper. Hanover, NH: Dartmouth University.
- Galinsky, Adam D., Joe C. Magee, Deborah H. Gruenfeld, Jennifer A. Whitson, and Katie A. Liljenquist. 2008. Power Reduces the Press of the Situation. *Journal of Personality and Social Psychology* 95 (6): 1450–66.
- Galtung, Johan. 1964. A Structural Theory of Aggression. *Journal of Peace Research* 1 (2):95–119.
- Gilpin, Robert. 1983. *War and Change in World Politics*. New York: Cambridge University Press.
- Glaser, Markus, Thomas Langer, and Martin Weber. 2005. Overconfidence of Professionals and Lay Men: Individual Differences Within and Between Tasks?. Working Paper. Available at <<https://ideas.repec.org/p/xrs/sfbmaa/05-25.html>>. Accessed 20 January 2015.
- Goldgeier, James M., and Philip E. Tetlock. 2001. Psychology and International Relations Theory. *Annual Review of Political Science* 4:67–92.
- Gordon, Michael E., L. Allen Slade, and Neal Schmitt. 1986. The “Science of the Sophomore” Revisited. *Academy of Management Review* 11 (1):191–207.
- Green, Donald P. 2009. Regression Adjustments to Experimental Data: Do David Freedman’s Concerns Apply to Political Science? Working Paper. New Haven, CT: Yale University.
- Gruenfeld, Deborah H., M. Ena Inesi, Joe C. Magee, and Adam D. Galinsky. 2008. Power and the Objectification of Social Targets. *Journal of Personality and Social Psychology* 95 (1):111–27.
- Guinote, Ana, Guillermo B. Willis, and Cristiana Martellotta. 2010. Social Power Increases Implicit Prejudice. *Journal of Experimental Social Psychology* 46 (2):299–307.
- Gunia, Brian C., Niro Sivanathan, and Adam D. Galinsky. 2009. Vicarious Entrapment. *Journal of Experimental Social Psychology* 45 (6):1238–44.
- Hafner-Burton, Emilie, Brad Leveck, David G. Victor, and James H. Fowler. 2012. A Behavioral Approach to International Legal Cooperation. Working Paper. Available at <http://papers.ssm.com/sol3/papers.cfm?abstract_id=1969905>. Accessed 17 January 2015.
- Hafner-Burton, Emilie M., D. Alex Hughes, and David G. Victor. 2013. The Cognitive Revolution and the Political Psychology of Elite Decision Making. *Perspectives on Politics* 11 (2):368–86.
- Hafner-Burton, Emilie M., and Alexander H. Montgomery. 2006. Power Positions International Organizations, Social Networks, and Conflict. *Journal of Conflict Resolution* 50 (1):3–27.

- Hatemi, Peter K., and Rose McDermott. 2012. Broadening Political Psychology. *Political Psychology* 33 (1):11–25.
- Heaven, Patrick C.L., Lee-Anne Organ, Sunila Supavadeeprasit, and Peter Leeson. 2006. A Study of Social Values, Right-wing Authoritarianism, and Social Dominance Orientation. *Personality and Individual Differences* 40 (3):599–608.
- Hertwig, Ralph, and Andreas Ortmann. 2003. Economists' and Psychologists' Experimental Practices. In *The Psychology of Economic Decisions*, edited by Isabelle Brocas and Juan Carrillo, 253–72. New York: Oxford University Press.
- 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):417–18.
- Horowitz, Michael, Rose McDermott, and Allan C. Stam. 2005. Leader Age, Regime Type, and Violent International Relations. *Journal of Conflict Resolution* 49 (5):661–85.
- Inesi, M. Ena. 2010. Power and Loss Aversion. *Organizational Behavior and Human Decision Processes* 112 (1):58–69.
- Kagan, Donald. 1995. *On the Origins of War and the Preservation of Peace*. New York: Random House.
- Kahneman, Daniel, and Jonathan Renshon. 2007. Why Hawks Win. *Foreign Policy* 158:34–38.
- Kahneman, Daniel, and Amos Tversky. 1979. Prospect Theory: An Analysis of Decision Under Risk. *Econometrica* 47 (2):263–91.
- . 1984. Choices, Values, and Frames. *American Psychologist* 39 (4):341–50.
- Kam, Cindy D. 2007. When Duty Calls, Do Citizens Answer? *Journal of Politics* 69 (1):17–29.
- Kelman, Herbert C. 1965. Social-psychological Approaches to the Study of International Relations: The Question of Relevance. In *International Behavior: A Social Psychological Analysis*, edited by Herbert C. Kelman, 565–607. New York: Holt, Rinehart, and Winston.
- Kertzer, Joshua D., and Kathleen M. McGraw. 2012. Testing the Microfoundations of Realism in Ordinary Citizens. *International Studies Quarterly* 56 (2):245–58.
- Kertzer, Joshua D., and Jonathan Renshon. 2014. Putting Things in Perspective: Mental Simulation in Experimental Political Science. Working Paper. Available at <<http://www.jonathanrenshon.net>>. Accessed 29 January 2015.
- Kindleberger, Charles P. 1981. Dominance and Leadership in the International Economy. *International Studies Quarterly* 25 (2):242–54.
- Larson, Deborah W., T.V. Paul, and William C. Wohlforth. 2014. Status and World Order. In *Status in World Politics*, edited by T.V. Paul, Deborah W. Larson, and William C. Wohlforth, 3–32. New York: Cambridge University Press.
- Larson, Deborah W., and Alexei Shevchenko. 2010. Status Seekers: Chinese and Russian Responses to US Primacy. *International Security* 34 (4):63–95.
- . 2014. Managing Rising Powers: The Role of Status Concerns. In *Status in World Politics*, edited by T.V. Paul, Deborah W. Larson, and William C. Wohlforth, 32–70. New York: Cambridge University Press.
- Lebow, Richard N. 2008. *A Cultural Theory of International Relations*. New York: Cambridge University Press.
- Lemke, Douglas, and William Reed. 1996. Regime Types and Status Quo Evaluations. *International Interactions* 22 (2):143–64.
- Lerner, Jennifer S., Deborah A. Small, and George Loewenstein. 2004. Heart Strings and Purse Strings. *Psychological Science* 15 (5):337–41.
- Lerner, Jennifer S., and Philip E. Tetlock. 1999. Accounting for the Effects of Accountability. *Psychological Bulletin* 125 (2):255–75.
- Levitt, Steven D., and John A. List. 2007. What Do Laboratory Experiments Measuring Social Preferences Reveal About the Real World? *Journal of Economic Perspectives* 21 (2):153–74.
- Levy, Jack S. 1987. Declining Power and the Preventive Motivation for War. *World Politics* 40 (1):82–107.
- Lopez, Anthony C., Rose McDermott, and Michael Bang Petersen. 2011. Evolution, Coalitional Psychology, and International Politics. *International Security* 36 (2):48–83.
- Magee, Joe C., Adam D. Galinsky, and Deborah H. Gruenfeld. 2007. Power, Propensity to Negotiate, and Moving First in Competitive Interactions. *Personality and Social Psychology Bulletin* 33 (2):200–12.

- Maner, Jon K., Matthew T. Gailliot, David A. Butz, and B. Michelle Peruche. 2007. Power, Risk, and the Status Quo. *Personality and Social Psychology Bulletin* 33 (4):451–62.
- Mani, Anandi, Sendhil Mullainathan, Eldar Shafir, and Jiaying Zhao. 2013. Poverty Impedes Cognitive Function. *Science* 341 (6149):976–80.
- Maoz, Zeev. 1989. Power, Capabilities, and Paradoxical Conflict Outcomes. *World Politics* 41 (2):239–66.
- . 2010. *Networks of Nations*. New York: Cambridge University Press.
- Markey, Daniel. 1999. Prestige and the Origins of War. *Security Studies* 8 (4):126–72.
- Marteau, Theresa M., and Hilary Bekker. 1992. The Development of a Six-item Short-form of the State Scale of the Spielberger State—Trait Anxiety Inventory. *British Journal of Clinical Psychology* 31 (3):301–6.
- McDermott, Rose. 2002a. Experimental Methodology in Political Science. *Political Analysis* 10 (4): 325–42.
- . 2002b. Experimental Methods in Political Science. *Annual Review of Political Science* 5:31–61.
- . 2011. Internal and External Validity. In *Handbook of Experimental Political Science*, edited by James N. Druckman, Donald P. Green, James H. Kuklinski, and Arthur Lupia, 27–41. New York: Cambridge University Press.
- McKeown, Timothy J. 1991. The Foreign Policy of a Declining Power. *International Organization* 45 (2): 257–79.
- McNemar, Quinn. 1946. Opinion-Attitude Methodology. *Psychological Bulletin* 43 (4):289–374.
- Meerts, Paul W. 2005. Entrapment in International Negotiations. In *Escalation and Negotiation in International Conflicts*, edited by I. William Zartman and Guy Olivier Faure, 111–41. New York: Cambridge University Press.
- Mercer, Jonathan. 1995. Anarchy and Identity. *International Organization* 49 (2):229–52.
- Midlarsky, Manus I. 1975. *Political Violence in the International System*. New York: Free Press.
- Miller, Neal E. 1941. The Frustration-Aggression Hypothesis. *Psychological Review* 48 (4):337–42.
- Mintz, Alex. 2004. Foreign Policy Decision Making in Familiar and Unfamiliar Settings. *Journal of Conflict Resolution* 48 (1):91–104.
- Mintz, Alex, Steven B. Redd, and Arnold Vedlitz. 2006. Can We Generalize from Student Experiments to the Real World in Political Science, Military Affairs, and IR? *Journal of Conflict Resolution* 50 (5):757–76.
- Morgenthau, Hans J. 1948. *Politics Among Nations*. New York: Alfred Knopf.
- Morton, Rebecca B., and Kenneth C. Williams. 2010. *Experimental Political Science and the Study of Causality: From Nature to the Lab*. New York: Cambridge University Press.
- O'Neill, Barry. 2001. *Honor, Symbols, and War*. Ann Arbor, MI: University of Michigan Press.
- Organski, Abramo F.K., and Jacek Kugler. 1980. *The War Ledger*. Chicago: University of Chicago Press.
- Pettit, Nathan C., Kevyn Yong, and Sandra E. Spataro. 2010. Reactions to the Prospect of Status Gains and Losses. *Journal of Experimental Social Psychology* 46 (2):396–401.
- Pu, Xiaoyu, and Randall L. Schweller. 2014. Status Signaling, Multiple Audiences, and China's Bluewater Naval Ambition. In *Status in World Politics*, edited by T.V. Paul, Deborah W. Larson, and William C. Wohlforth, 141–62. New York: Cambridge University Press.
- Renshon, Jonathan. Forthcoming. Status Deficits and War. *International Organization*.
- Renshon, Jonathan, Joa Julia Lee, and Dustin Tingley. 2014. Physiological Arousal and Political Beliefs. *Political Psychology*. DOI: 10.1111/pops.12173
- Renshon, Jonathan, and Stanley A. Renshon. 2008. The Theory and Practice of Foreign Policy Decision Making. *Political Psychology* 29 (4):509–36.
- Rhamey, J. Patrick, and Bryan R. Early. 2013. Status-Seeking Behavior and Olympic Performance. *International Area Studies Review* 16 (3):244–61.
- Ronay, Richard, and William Von Hippel. 2010. Power, Testosterone, and Risk-taking. *Journal of Behavioral Decision Making* 23 (5):473–82.
- Rose, Gideon. 1998. Neoclassical Realism and Theories of Foreign Policy. *World Politics* 51 (1):144–72.
- Rudman, Laurie A., Corinne A. Moss-Racusin, Julie E. Phelan, and Sanne Nauts. 2011. Status Incongruity and Backlash Effects. *Journal of Experimental Social Psychology* 48 (1):165–79.

- Scheepers, Daan. 2009. Turning Social Identity Threat into Challenge. *Journal of Experimental Social Psychology* 45 (1):228–33.
- Sears, David O. 1986. College Sophomores in the Laboratory. *Journal of Personality and Social Psychology* 51 (3):515–30.
- See, Kelly E., Elizabeth W. Morrison, Naomi B. Rothman, and Jack B. Soll. 2011. The Detrimental Effects of Power on Confidence, Advice Taking, and Accuracy. *Organizational Behavior and Human Decision Processes* 116 (2):272–85.
- Sherman, Gary D., Jennifer Lerner, Robert Josephs, Jonathan Renshon, and James Gross. 2013. Who Will Lead? Testosterone, Cortisol and Elite Leaders. Working Paper.
- Sherman, Gary D., Julia J. Lee, Amy J.C. Cuddy, Jonathan Renshon, Chris Oveis, James J. Gross, and Jennifer S. Lerner. 2012. Leadership Is Associated with Lower Levels of Stress. *Proceedings of the National Academy of Sciences* 109 (44):17903–7.
- Simonson, Itamar, and Peter Nye. 1992. The Effect of Accountability on Susceptibility to Decision Errors. *Organizational Behavior and Human Decision Processes* 51 (3):416–46.
- Singer, J. David. 1961. The Level-of-Analysis Problem in International Relations. *World Politics* 14 (1): 77–92.
- . 1988. Reconstructing the Correlates of War Dataset on Material Capabilities of states, 1816–1985. *International Interactions* 14 (2):115–32.
- Singer, J. David, and Melvin Small. 1966. The Composition and Status Ordering of the International System: 1815–1940. *World Politics* 18 (2):236–82.
- Siverson, Randolph M. 1995. Democracies and War Participation. *European Journal of International Relations* 1 (4):481–89.
- Smith, Pamela K., Nils B. Jostmann, Adam D. Galinsky, and Wilco W. van Dijk. 2008. Lacking Power Impairs Executive Functions. *Psychological Science* 19 (5):441–47.
- Smith, Pamela K., Daniël H. Wigboldus, and Ap Dijksterhuis. 2008. Abstract Thinking Increases One's Sense of Power. *Journal of Experimental Social Psychology* 44 (2):378–85.
- Staw, Barry M. 1981. The Escalation of Commitment to a Course of Action. *Academy of Management Review* 6 (4):577–87.
- Staw, Barry M., and Jerry Ross. 1989. Understanding Behavior in Escalation Situations. *Science* 246 (4927):216.
- Taliaferro, Jeffrey W. 1998. Quagmires in the Periphery. *Security Studies* 7 (3):94–144.
- Tetlock, Philip E. 1985. Accountability: The Neglected Social Context of Judgment and Choice. *Research in Organizational Behavior* 7:297–332.
- Thaler, Richard. 1980. Toward a Positive Theory of Consumer Choice. *Journal of Economic Behavior and Organization* 1 (1):39–60.
- Tingley, Dustin H., Julia J. Lee, and Jonathan Renshon. 2014. Physiological Responses to Shifting Bargaining Power: Micro-foundations of Commitment Problems. Working Paper. Available at <<http://www.jonathanrenshon.net>>. Accessed 29 January 2015.
- Tingley, Dustin H., and Barbara Walter. 2011. Reputation Building in International Relations. *International Organization* 65 (2):343–65.
- Tomz, Michael. 2007. Domestic Audience Costs in International Relations. *International Organization* 61 (4):821–40.
- Tomz, Michael, Jason Wittenberg, and Gary King. 2003. Clarify: Software for Interpreting and Presenting Statistical Results. *Journal of Statistical Software* 8 (1):1–30.
- Turner, John C., Rupert J. Brown, and Henri Tajfel. 1979. Social Comparison and Group Interest in Ingroup Favouritism. *European Journal of Social Psychology* 9 (2):187–204.
- Tyszka, Tadeusz, and Piotr Zielonka. 2002. Expert Judgments: Financial Analysts Versus Weather Forecasters. *Journal of Psychology and Financial Markets* 3 (3):152–60.
- Verba, Sidney. 1964. Simulation, Reality, and Theory in International Relations. *World Politics* 16 (3): 490–519.
- Volgy, Thomas J., Renato Corbetta, J. Patrick Rhamey, Ryan G. Baird, and Keith A. Grant. 2014. Status Considerations in International Politics and the Rise of Regional Powers. In *Status in World Politics*,

- edited by T.V. Paul, Deborah W. Larson, and William C. Wohlforth, 141–62. New York: Cambridge University Press.
- Volgy, Thomas J., and Stacey Mayhall. 1995. Status Inconsistency and International War. *International Studies Quarterly* 39 (1):67–84.
- Von Hippel, Karin. 1996. Sunk in the Sahara. *Journal of North African Studies* 1 (1):95–116.
- Wallace, Michael. 1973. *War and Rank Among Nations*. Lexington, MA: Lexington Books.
- Walt, Stephen. 2010. Cutting Losses in Wars of Choice. In *The Prudent Use of Power in American National Security Strategy*, edited by Stephen Van Evera and Sidharth Shah, 131–56. Cambridge, MA: Tobin Project.
- Watson, David, Lee A. Clark, and Auke Tellegen. 1988. Development and Validation of Brief Measures of Positive and Negative Affect. *Journal of Personality and Social Psychology* 54 (6):1063–70.
- Wechsler, David. 1997. *Wechsler Adult Intelligence Scale*. San Antonio, TX: Psychological Corporation.
- Welch, David A. 2005. *A Theory of Foreign Policy Change*. Princeton, NJ: Princeton University Press.
- Wohlforth, William C. 1993. *Power and Perceptions During the Cold War*. Ithaca, NY: Cornell University Press.
- . 2009. Unipolarity, Status Competition, and Great Power War. *World Politics* 61 (1):28–57.
- Wolf, Reinhard. 2011a. Recognition and Disrespect Between Persons and Peoples. In *The International Politics of Recognition*, edited by Thomas Lindemann and Erik Ringmar, 39–57. Boulder, CO: Paradigm.
- . 2011b. Respect and Disrespect in International Politics. *International Theory* 3 (1):105–42.
- Wong, Kin Fai Ellick, Michelle Yik, and Jessica Y.Y. Kwong. 2006. Understanding the Emotional Aspects of Escalation of Commitment. *Journal of Applied Psychology* 91 (2):282–97.
- Wood, Steve. 2013. Prestige in World Politics. *International Politics* 50 (3):387–411.
- Zink, Caroline F., Yunxia Tong, Qiang Chen, Danielle S. Bassett, Jason L. Stein, and Andreas Meyer-Lindenberg. 2008. Neural Processing of Social Hierarchy in Humans. *Neuron* 58 (2):273–83.