Resolve, Time, and Risk

Joshua D. Kertzer

Abstract Why do some actors in international politics display remarkable persistence in wartime, while others "cut and run" at the first sign of trouble? I offer a behavioral theory of resolve, suggesting that variation in time and risk preferences can help explain why some actors display more resolve than others. I test the theory experimentally in the context of public opinion about military interventions. The results not only help explain why certain types of costs of war loom larger for certain types of actors but also shed light on some of the behavioral revolution's contributions more broadly.

Why do some leaders and publics display remarkable persistence in military interventions—the British in the Malayan emergency, for example—while others cut and run, like the US in Somalia in 1993? When political scientists ask this question, the answer we typically give has to do with resolve. Resolve is a ubiquitous ingredient in International Relations (IR) theory, used to explain everything from why states win wars, to how they prevail during crisis bargaining, thereby preventing conflict from breaking out.[1] Rationalist approaches to the study of international conflict revolve around resolve: if both sides are aware of each other's levels of motivation, the less resolute side backs down before a crisis can even take place.[2] Likewise, it is motivation, not muscle, that is frequently used to explain why great powers fare so poorly in asymmetric conflict, and why the United States was never able to push the North Vietnamese to their "breaking point" during the Vietnam War.[3] Yet although we often explain behavior and outcomes in IR by pointing to variation in actors' levels of resolve, we have trouble modeling this variation itself, and explaining why some actors are more resolved than others.

I address this challenge by offering a behavioral theory of resolve, suggesting that variation in time and risk preferences can help explain why some actors display more resolve than others. I seek to explain variations in resolve not only because of the phenomenon's clear scholarly and political importance but also because it serves as an excellent conceptual vehicle for exploring the new behavioral revolution in IR.

Previous versions of this project benefited from remarkably helpful feedback from audiences at Binghamton, Dartmouth, Harvard, Illinois, Kansas, McGill, NYU, OSU, Princeton, UNC, both USCs, UCLA, UC San Diego, UC Santa Barbara, Yale, ISA, and Peace Science, and too many people to name. Portions of the article draw on material from *Resolve in International Politics* by Joshua D. Kertzer. Copyright © 2016 by Princeton University Press. Reprinted by permission.

1. For example, Maoz 1983; Schelling 1966; Snyder and Diesing 1977.
2. Fearon 1995; Powell 1987.
3. Mack 1974; Mueller 1980.

International Organization 71, Supplement 2017, pp. S109–S136
© The IO Foundation, 2017.
doi:10.1017/S0020818316000424

https://doi.org/10.1017/S0020818316000424 Published online by Cambridge University Press

Like the structural realists they sought to supplant, the rationalist approaches that have been prominent in mainstream IR over the past several decades tend to privilege structure over agency, based on a "methodological bet" that greater theoretical progress can be made in International Relations by focusing on environmental features rather than actor-level characteristics.[4] Although this tendency has fostered many rich and remarkable contributions, it has also stymied our attempts to understand resolve, frequently reducing it to an actor's cost of war, and thereby leading to a number of both empirical and conceptual conundrums. My claim is not just that we ignore important individual-level variation when we reduce resolve to a situational feature, but that if resolve is defined by *resisting* situationally induced pressures, it is not clear that studying resolve situationally is in fact studying resolve at all.

Building on behavioral research from elsewhere in the social sciences, I show how we can explain variation in resolve by pointing to variation in actors' dispositional characteristics, which can open up the black box of "costliness" and explain why certain types of costs of war loom larger for certain types of actors. Importantly, I focus on two actor-level traits— time and risk preferences—that play important supporting roles in rationalist models of conflict. In this sense, consistent with Rathbun, Kertzer, and Paradis's contribution in this issue, I show that behavioral approaches are by no means incompatible with rationalist frameworks, and that acknowledging the many ways actors differ from one another does not preclude the possibility of explaining behavior systematically.[5] What is particularly valuable about behavioral approaches, however, is their emphasis on acknowledging systematic differences in how actors perceive the situations they face, rather than assuming agreement and taking perceptions for granted. Exploring resolve experimentally in the context of public opinion about military interventions, I show how more patient individuals display more resolve, and are less sensitive to casualties, while individuals with more extreme risk preferences tend to display more resolve, with more risk-averse individuals exhibiting greater sensitivity both to the human costs of fighting, and to the reputational costs of backing down. Explaining variation in resolve thus requires exploring variation across actors themselves, rather than just the environments they face.

The discussion that follows has four parts. I begin by defining resolve and reviewing the manner in which it has been used in IR, suggesting that many of the difficulties we have encountered in explaining resolve come from trying to shoehorn the concept into a situational straitjacket. Building from work in the "behavioral revolution" taking place throughout the social sciences, I argue that linking resolve to specific dispositional characteristics provides more satisfactory microfoundations, and present a series of hypotheses on the relationship between resolve, time, and risk. Third, I turn to experimental methods to test my theoretical framework in a public-opinion context, presenting the results from a novel two-stage laboratory experiment

4. See Hafner-Burton et al. 2017.
5. Rathbun, Kertzer, and Paradis 2017.

https://doi.org/10.1017/S0020818316000424 Published online by Cambridge University Press

that models both the selection into and duration of support for military interventions, manipulating situational features of the military intervention while measuring dispositional variables using techniques developed by behavioral economists. I conclude by briefly discussing the ramifications of the findings, for both theories of resolve, and the behavioral revolution in IR more broadly.

Resolve in International Politics

Following the Oxford English Dictionary, I define resolve as "firmness or steadfastness of purpose," maintaining a policy despite contrary inclinations or temptations to back down. In its emphasis on determined, sustained effort despite temptations to the contrary, resolve can therefore be considered synonymous with willpower, as well as related forms of self-regulation like self-control (the deliberate use of willpower to avoid undesirable actions), and the opposite of weakness of will and other such antonyms. Defining resolve in this manner has two advantages. First, it reflects the way we use these words in ordinary language. To describe an actor as "resolute" intuitively means the same thing as to argue the actor is persistent, motivated, committed, determined, and so on. Second, it allows the analysis that follows to be grounded in an interdisciplinary fashion. The past several decades have witnessed a *Risorgimento* of resolve across the social sciences.[6] This growth has occurred in tandem despite each literature retaining its own nomenclature and applying the concept to different domains: political scientists restrict resolve to military contexts by defining the term as "willingness to fight," while economists tend to define self-control purely in terms of consumption decisions, and sociologists focus on willpower in the realm of refraining from socially undesirable actions. Despite these differences, the underlying construct remains the same: a resolute actor is engaged in "an intensity of feeling": a firmness or steadfastness of purpose, resisting temptations to the contrary.[7] In this sense, resolve is a *second-order* phenomenon, in that it refers not to the substance or content of an actor's desire—whether to fight, quit smoking, save money, or so on—but to the steadfastness, dogged persistence, or "stick-to-itiveness" with which it is being pursued.[8]

Resolve thus defined is a frequent protagonist in IR theory, across levels of analysis. At the state or leader level, resolve looms large in multiple bodies of scholarship, ranging from deterrence theory with its emphasis on forming reputations for resolve, to the crisis bargaining literature more broadly, which argues that states embroiled in crisis negotiations are partaking in games of risk that are won not by strength, but by

6. For example, Baumeister, Vohs, and Tice 2007; Bénabou and Tirole 2004; Gottfredson and Hirschi 1990; Thaler and Shefrin 1981.

7. Young 1968, 33.

8. Frankfurt 1971. The philosophy of action literature has traditionally explored resolve in the context of upholding better judgments but, if resolve merely refers to a firmness of purpose irrespective of its content, one can be as resolved to indulge as one is to abstain.

https://doi.org/10.1017/S0020818316000424 Published online by Cambridge University Press

"nerve."[9] International security scholars also link states' resolve to their military outcomes. In the shadow of the Vietnam War, for example, Rosen claimed that to understand which side wins the war, we need to pay attention to each party's "willingness to suffer," and Mueller pointed to the importance of the "breaking point" in determining who accepts defeat.[10] Just as crisis bargaining theorists argue that war can be prevented if the parties recognize the true balance of motivation, scholars of war termination suggest that wars end once the combatants learn each other's levels of resolve.[11]

IR scholars are also interested in resolve at the level of the mass public, which is especially important for our purposes, given the empirical context I investigate. Although some rationalist theories suggest democracies are often able to avoid war because they are better at signaling their resolve, scholars of counterinsurgencies and asymmetric conflict frequently inquire whether democratic publics are too "soft" or cost intolerant to be resolute.[12] Even if the public is not casualty phobic, "rally around the flag" effects end just as quickly as they begin, and support for conflict decays drastically over time.[13] Similarly, when classical IR theorists spoke of "national morale" or the "national will" as sources of power, they were explicitly assuming that the resolve of the public mattered for issues of war and peace, as do contemporary commentators who lament, as Lyndon Johnson did, that America's wars are won abroad but lost at home.[14]

Searching for Microfoundations

References to resolve permeate the international security literature. Yet although resolve is central to our understanding of how international security works, there are at least two things we do not understand about resolve. First, as Sartori notes, the term has been used in a variety of different ways by different scholars so that conceptually it is often unclear what resolve actually is.[15] IR scholars often suggest that states have three pieces of private information: capabilities, intentions, and resolve, yet resolve is frequently presented in a manner that renders it analytically indistinct from the other two. Some approaches lump in resolve with capabilities, as in classic IR theorists who wrote about "national morale" or "national will" as a component of power.[16] Others reverse the causal arrow and see capabilities as a source

9. Jervis 1976, 1979; Powell 2003.
10. Mueller 1980; Rosen 1972.
11. Goemans 2000; Ramsay 2008.
12. Gelpi, Feaver, and Reifler 2009; Lyall 2010; Mack 1974; Mueller 1971; Schultz 1999.
13. Baum 2002; Reiter and Stam 2002, chapter 7.
14. Morgenthau 1985; Wright 1964.
15. Sartori 2005, 7.
16. Cline 1975; Morgenthau 1985; Organski 1968; Wright 1964.

https://doi.org/10.1017/S0020818316000424 Published online by Cambridge University Press

of resolve,[17] or use the terms indistinguishably.[18] For others, resolve is used synonymously with aims or intentions.[19] Indeed, uncertainty about an actor's resolve is often framed as ambiguous or misperceived intentions: whether Berlin knew if St. Petersburg wanted war or not in 1914, the veracity of Saddam Hussein's assumption that the United States would stand idly by during the 1990 Iraqi invasion of Kuwait, the US trying to convince Argentina that Great Britain's threats of war in the Falklands Crisis were not bluffs, and so on.[20]

The issue here goes beyond the usual slings and arrows of polysemy. Although some conceptual ambiguity is likely inevitable when employing a concept as ubiquitous as resolve, many of these quandaries also stem from the absence of microfoundations. Despite its prominent explanatory role in our theories, we do not have a clear sense of where resolve comes from, and what the individual-level "cogs and wheels" might be that act as mechanisms linking cause and effect.[21] This absence of well-specified microfoundations is consequential because it tempts us to tautologically infer resolve from the same outcomes we use it to explain. Policy-makers are particularly prone to this—when the Islamic State regained control of Ramadi in May 2015 despite being outnumbered by Iraqi Security Forces, Secretary of Defense Ashton Carter attributed the Iraqi defeat to an absence of resolve, a diagnosis auspiciously absent *ex ante*—but similar maneuvers are made by political scientists as well. Studying resolve in this manner problematically turns the concept into a residual category used after the fact to explain outcomes we were unable to explain beforehand, making it difficult to subject our theories of resolve to rigorous empirical testing.

The mainstream IR literature on resolve—particularly in the rational choice tradition—has made two analytic moves that have exacerbated these conceptual quandaries, and made the search for microfoundations difficult. First, consistent with rationalism's "methodological bet" that greater theoretical progress can be made in IR by focusing on structural features than actor-level heterogeneity, much of IR has studied resolve *situationally*, treating it primarily as a feature of the situation actors face, rather than characteristics of actors themselves.[22] Similar to classic works in psychology that explored how people and pigeons alike respond predictably to external stimuli and reinforcements, much of the IR literature has defined resolve in terms of costs and benefits, especially in the crisis bargaining literature: an actor will be resolved in a crisis bargaining situation when the stakes are high relative to the costs, and irresolute when they are low, such that actors are systematically more resolute about some issues than about others.[23]

17. For example, Meirowitz and Sartori 2008, 329n4; Snyder and Diesing 1977.
18. Slantchev 2010; Smith and Stam 2004, 808.
19. Fearon 1994; Jervis 1976; Trager 2011.
20. Copeland 2000, 90–91; Kydd 2003; Mearsheimer 2001, 38.
21. Elster 1989, 3. On the importance of microfoundations in IR, see Kertzer forthcoming.
22. On the rationalist-strategic methodological bet, see Hafner-Burton et al., 2017.
23. Herrnstein 1970; Mischel and Grusec 1967.

https://doi.org/10.1017/S0020818316000424 Published online by Cambridge University Press

There are a variety of conceptual distinctions here: some approaches fully reduce the "balance of resolve" to the "balance of interests" and thereby treat an actor's resolve as its cost of conflict, while others treat an actor's costs as one of several determinants of resolve.[24] Moreover, reducing resolve to cost-benefit calculations belies the question both of *which* costs of war matter—the human costs of fighting, the reputational costs of defeat and withdrawal, and so on—as well as *whose* costs matter, since the cost of war is typically unevenly distributed, borne more heavily by some constituencies than others.[25] Although hardly a unified voice, then, what these situational theories all share is an automaticity through which costs translate into action: sufficiently raise the costs, and all actors will back down, irrespective of any internal factor that affects how they respond to these external stimuli.[26]

Although succinct and tractable, purely situational theories of resolve ultimately make for dissatisfying microfoundations, since in treating resolve as a property of the environment rather than the actor, they deliberately bracket micro-level processes altogether. Indeed, as Hafner-Burton and colleagues point out, these approaches purposely sidestep the question of actor-level heterogeneity, which is seen as being worthy of exploration only as a matter of last resort.[27] In the case of resolve, this raises concerns about construct validity, since if we think of resolve as maintaining a policy *despite* contrary inclinations or temptations to back down, it implies that resolve involves resisting situationally induced pressures to retreat or reverse course; to reduce resolve to utility or a cost-benefit calculation purges it of its second-order quality. Moreover, when we talk about someone being resolved, we are usually referring to something that emanates from within rather than merely dictated from without—a property of the actor itself rather than a feature of the situation. Military thinkers like Napoleon and Foch contrasted resolve with material factors out of the assumption that resolve involves being unmoved by material disadvantages. Our lay theories of resolve, then, are *dispositional* in nature. As long as resolve is one of the first explanations we offer for our outcomes of interest, we cannot reluctantly turn to actor-level heterogeneity only as a matter of last resort.

Thus, a second body of work in IR has acknowledged a dispositional component to resolve, similar to research in social psychology treating willpower as a trait or individual difference.[28] Sartori, for example, notes that although some deterrence theorists have viewed resolve solely as a function of the issues at stake in a particular crisis, others "maintain that it is an enduring, dispositional quality, that some states generally are more willing to fight than others."[29] Both formal and empirical IR scholarship note that actors vary in their sensitivity to costs, or "willingness to

24. Kydd 2003; Mattes and Morgan 2004; Morrow 1987; Pape 1996.
25. Chiozza and Goemans 2004.
26. Pape 1996; Sullivan 2007.
27. See Hafner-Burton et al. 2017.
28. Hagger et al. 2010.
29. Sartori 2005, 45.

https://doi.org/10.1017/S0020818316000424 Published online by Cambridge University Press

pay."[30] Similarly, two of the dominant research traditions in the democracies in war literature—the "selection effects" work that argues that democracies are more likely to win their wars because they are more cautious about which wars are worth fighting, and the literature investigating whether the advantages of democracy in battle decline over time because democratic publics are too casualty-shy to sustain lengthy combat operations—both acknowledge the importance not just of costs, but of cost sensitivity.[31] In much of the literature, the issue is not whether democracies have higher costs of war than autocracies, but how tolerant they are of the costs they do face.

Although cost sensitivity brings us closer to microfoundations than purely situational theories of resolve do, it still makes for a dissatisfying explanation for resolve in that it suffers from "the problem of interiority."[32] That is, it explains resolve as a result of actors being less sensitive to the cost of war. This merely pushes the causal story downwards, leaving unanswered the question of why some actors are more sensitive to the costs of war in the first place—akin to treating heterogeneity as an empirical nuisance to be controlled for with fixed effects, rather than something meriting theoretical investigation in its own right. Thus, the task I undertake here is to go beyond cost sensitivity and point to specific actor-level characteristics to investigate where cost sensitivity comes from.

Time and Risk Preferences

In this paper, I explore the roles of two actor-level characteristics: time preferences and risk preferences. The attraction of these actor-level traits is threefold. First, they are routinely discussed as predictors of resolve, willpower, or self-control in social psychology and behavioral economics. Following Rosenau, if theorizing is about asking "of what is this an instance?" then turning to these traits allows us to link research on resolve in IR with work on resolve more generally.[33]

Second, one of the central empirical challenges of the behavioral revolution is in specifying how actors perceive and construct the situations they face.[34] Time and risk preferences are two variables that help address this challenge with situational theories of resolve by opening up the black box of "costliness" and predicting why certain types of actors are more or less sensitive to certain types of costs of war. This, I believe, is the major contribution that the "behavioral revolution" can make to International Relations. As Simon argued thirty years ago, theories of rational utility maximizing are only as useful as the auxiliary assumptions they rely on about where actors derive their utility from. Although situational theories of

30. Jervis 1976; Mueller 1980; Slantchev 2003; Stam 1996.
31. Gelpi, Feaver, and Reifler 2009; Sullivan 2007; Valentino, Huth, and Croco 2010.
32. Reus-Smit 2003.
33. Rosenau 1980.
34. Hafner-Burton et al. 2017; Herrmann 1988.

https://doi.org/10.1017/S0020818316000424 Published online by Cambridge University Press

resolve appear concise and clear in their predictions, some of this clarity comes from a sort of "iceberg parsimony," where much of the theory's mass is hidden under water, and it is this subaquatic portion that is doing much of the work.[35] It is one thing to say that actors will be resolved when their payoff structures tell them to be, but it is another to specify where these payoffs come from in the first place.

Finally, unlike some other candidates from elsewhere in the social sciences, time and risk preferences are helpful building blocks because they are already familiar quantities to political scientists and play important roles in rationalist models of conflict. In this sense, if this special issue's purpose is to demonstrate the value of building more realistic assumptions about actors' preferences, beliefs, and decision-making processes into our theories of world politics, one of my goals here is to show that some of this infrastructure is currently lying dormant within our extant theoretical models, and that the gulf between psychological and rational theories of world politics is thus not as wide as is often stated.[36]

Time Preferences

Although there has been growing interest in time horizons in IR, the term has been used to refer to both the length of the time window in which leaders must act and the notion of *time preferences*, a variable measuring the extent to which actors value the future compared to the present.[37]

A theory of resolve based on time preferences is in many ways intuitive, since the two factors are frequently conceptualized in relation to one another. We inherently think of someone who is resolved as being "in it for the long haul," and define a lack of willpower as "excessive preference for the present," or as a faculty that allows us to pursue our ends over time.[38] Similarly, one of the most popular ways of modeling willpower or self-control problems in the behavioral sciences is through "dual-self" models in which choice decisions are made as the result of a bargaining process between a present- and a future-oriented self, buttressed by work in neuroscience suggesting that immediate gratification and long-term considerations activate different parts of the brain.[39] For these reasons, I argue that longer time horizons are associated with greater resolve. Suicide bombers, for example, are usually deemed resolute precisely because they have long time horizons: carrying out an attack requires sacrificing oneself in the present either to confer benefits on future generations or to enjoy rewards promised in eternity.[40]

35. Simon 1985. This charge is similar to the constructivist critique of structural realism for basing its explanatory power on implicit assumptions about the "distribution of interests." Wendt 1999, chapter 3.
36. McDermott 2004; Mercer 2005; Rathbun, Kertzer, and Paradis 2017.
37. Krebs and Rapport 2012; Streich and Levy 2007; Toft 2006.
38. Bénabou and Tirole 2002, 879; Velleman 2007.
39. Bénabou and Tirole 2004; Benhabib and Bisin 2005; Elster 1986.
40. Toft 2006.

https://doi.org/10.1017/S0020818316000424 Published online by Cambridge University Press

H1: Actors with longer time horizons (more patient time preferences) will display more resolve.

We can also think of time preferences in an interactionist framework since an actor's time preferences affect its sensitivity to costs and benefits that occur in the future relative to those that occur in the present. Individuals typically face costs and benefits occurring at multiple points in time. Just as choosing between carrots or cheesecake, for example, requires managing tradeoffs between present and future costs and benefits (cheesecake provides immediate gratification but negative long-term consequences, while carrots typically provide fewer benefits in the short term but greater benefits in the long term), choosing between continuing to prosecute a war or to terminate it short of victory involves a similar set of temporally arranged tradeoffs. The nature of the intertemporal tradeoff varies across issues or domains but many military interventions are characterized by a temporal configuration in which the human costs of war are paid up front, whereas reputational benefits are incurred in the future. Indeed, as the literature on cooperation under anarchy suggests, the more actors discount the future, the less reputational costs matter.[41] Thus, once actors with longer time horizons enter into a war, they should outlast those with shorter time horizons since they should be more willing to pay the short-term human costs of war in exchange for the longer-term benefits victory is posited to bring.

H2: Actors with longer time horizons (more patient time preferences) will be relatively less sensitive to casualties.

Risk Preferences

In IR, we frequently think of risk attitudes as a source of resolve and the most influential work on coercion in international politics calls coercive bargaining "competition in risk-taking."[42] Given the literature equating resolve with an actor's level of "critical risk," the greatest risk of disaster a state will tolerate to prevail in a crisis, one might expect resolve to increase with risk acceptance: the more tolerant actors are of risks, the more resolute they should be.[43] Yet given the insights of the behavioral decision-making literature, there are two reasons one might be skeptical of this hypothesized relationship. First, although IR scholars tend to see risk-acceptant behavior as an indication of a high level of resolve, sociologists have typically understood risk-seeking behavior as an indication of a lack of self-control; if resolve is indeed related to willpower and self-control, this contradiction is some cause for concern.[44] Second, and more importantly, whereas both rationalists and

41. Axelrod 1984; Oye 1985.
42. Morrow 1987; Schelling 1966.
43. Horowitz and Stam 2014; McDermott 1998.
44. For example, Gottfredson and Hirschi 1990.

https://doi.org/10.1017/S0020818316000424 Published online by Cambridge University Press

constructivists frequently rely on assumptions of common knowledge and intersubjectivity, one of the key insights of the behavioral revolution has been to emphasize the importance of modeling how actors perceive the situations they face, rather than assuming agreement and taking perceptions for granted.[45] Similar to Knight's distinction between risk and uncertainty, much of the political domain is not characterized by the kind of objective, measurable probabilities that would facilitate a clean mapping between risk attitudes and specific policy choices, without also taking actors' *risk perceptions* into account.[46] Many public debates, whether on climate change or counterterrorism, center not on whether we should accept risk or not, but rather, on contesting which choices count as risky in the first place.

This dynamic is especially likely with military interventions. IR scholars who have thought about risk have often done so in a dovish way, associating risk with fighting. This is consistent with the liberal way we as a discipline tend to think about war in general, whether in the form of realists focusing on exculpatory causes of war like the security dilemma, or rationalists representing war as a form of market failure. Yet, just as actors vary in their causal beliefs about threats, both fighting *and* backing down can be perceived as risky.[47] Military interventions are typically characterized by ill-structured decision environments, with little intersubjective agreement about what the relevant probabilities are. As the civil war in Iraq dragged on in 2004, for example, neoconservatives argued that the reputational consequences of withdrawal made "cutting and running" riskier than pressing onward, whereas opponents of the war suggested that sinking into a quagmire was far riskier than a quick exit. If actors facing the same situation nonetheless have contradictory risk perceptions, the relationship between risk and resolve becomes more complex than our standard models assume.

We can formalize the relationship between an actor's risk attitudes, risk perceptions, and resolve in military interventions in a simple stylized model:

$$y = (b - f)\rho + \varepsilon$$

in which y, resolve, is a function of the product of an individual's risk preferences (ρ, where positive values indicate greater degrees of risk aversion) and the difference between the perceived riskiness of fighting (f) versus backing down (b), across n individuals. For heuristic purposes, suppose four ideal types of actors, each of whom represents a different combination of risk perceptions: a *dovish* profile in which fighting is seen as risky and backing down is not ($b = 0$, $f = 1$), a *hawkish* profile in which backing down is seen as risky and fighting is not ($b = 1$, $f = 0$), an *ambivalent* profile in which both fighting and backing down are seen as risky ($b = 1$, $f = 1$), and a *neutral* profile in which neither fighting nor backing down are seen as risky ($b = 0$, $f = 0$).

45. Fearon and Wendt 2002; Grynaviski 2014; Herrmann 1988.
46. Knight 1921; Slovic 1987. For an IR application of this distinction, see Nelson and Katzenstein 2014.
47. Kertzer and Brutger 2016; Saunders 2011.

https://doi.org/10.1017/S0020818316000424 Published online by Cambridge University Press

For doves, because pressing on in a quagmire is seen as risky, risk aversion will be associated with decreased resolve—the traditional way we understand risk preferences in IR. For hawks, because backing down in the face of reputational consequences is seen as risky, risk aversion will be associated with *increased* resolve—the opposite of the conventional understanding of risk in the IR literature.[48] On average, then, the relationship between risk orientations and resolve in any given sample thus depends on two factors that are of particular theoretical interest. First is the distribution of risk perceptions, both in terms of $\frac{n_{b=f}}{n}$, the proportion of ambivalent and neutral types in the sample (increases in which attenuate the relationship between risk preferences and resolve), and $\frac{n_{b=1,f=0}}{n_{b=0,f=1}}$, the ratio of doves to hawks (which affects the slope of the effect). In general, given motivated reasoning and the difficulties people have in grappling with tradeoffs, it is plausible to assume that $cor(b, f) < 0$, such that believing that fighting is risky decreases the probability of perceiving backing down as risky.[49] In this case, the proportion of ambivalent and neutral types in the sample will be relatively low and our chief interest in the distribution of risk preferences concerns the ratio of doves to hawks.

Second is $cor(b - f, \rho)$, the correlation between risk perceptions and risk preferences. If risk preferences and risk perceptions are uncorrelated (such that actors' general comfort toward risk is independent of what they believe to be risky), on average we would expect risk preferences to have a linear relationship with resolve, the precise slope of which will depend on the ratio of doves to hawks. However, if risk preferences and risk perceptions are correlated in a given context, the ratio of doves to hawks will shift at different levels of risk aversion, producing a nonlinear relationship between risk preferences and resolve. If hawks in the sample are generally more risk averse than doves—consistent with conservative foreign policy preferences being linked to decreased tolerance of uncertainty and ambiguity, and lower levels of openness to experience—the relationship between risk and resolve will resemble a bell curve, with relatively risk-averse and relatively risk-acceptant actors displaying more resolve than participants with moderate levels of risk aversion; those with less extreme risk preferences will be more likely to cut and run.[50]

H3a: Risk aversion has a nonlinear relationship with resolve.

Finally, one of the implications of thinking about these two countervailing sets of risk preferences is that risk aversion should moderate the impact of both the costs of

48. For ambivalent and neutral types, there should be no relationship between risk orientation and resolve, since the two countervailing risk perceptions should cancel one another out.

49. Herrmann 2017; Steinbruner 1974; Tetlock 2003.

50. Jost et al. 2003. An alternative way of exploring the relationship would involve obtaining direct measures of actors' risk perceptions, whereupon we would expect simpler linear effects of risk orientations conditional on risk perceptions.

https://doi.org/10.1017/S0020818316000424 Published online by Cambridge University Press

fighting, and the costs of backing down, increasing sensitivity to both: in an experimental context, on average, the effects of casualty treatment T_c should increase with risk aversion such that more risk-averse participants should be more sensitive to casualties ($y = (b - T_c f)\rho + \varepsilon, \frac{dy}{dT_c} = -f\rho$), while the effects of reputation cost treatment T_r should decrease with risk aversion, such that more risk-averse participants should be more sensitive to reputation costs ($y = (T_r\, b - f)\rho + \varepsilon, \frac{dy}{dT_r} = b\rho$).

H3b: Risk aversion increases sensitivity to both the costs of fighting and the costs of backing down.

Research Design and Materials

I test these hypotheses experimentally in the context of public opinion about military interventions, one of the defining issues of the past decade, with questions about the sustainability of the US missions in Afghanistan and Iraq occupying a prominent position in most Western countries' foreign policy agendas. I employ a public opinion experiment for four reasons. First, the question of the determinants of the public's resolve—what classical IR scholars used to call "national will" or "national morale"—is of both theoretical and empirical importance in an era when lingering questions remain about publics' willingness to shoulder military obligations overseas, and their tendency to succumb to Morgenthau's concerns about the public's "paralysis of will."[51] Second, studying resolve in a public opinion context allows us to harness the power of experimental methods, which are uniquely well suited to testing the individual-level microfoundations implicated in our theories, as work in this special issue on commitment problems, compliance with international law, responses to rising power, threat perception, and support for free trade makes clear.[52]

Third, experimental methods also offer important measurement advantages: resolve is notoriously difficult to study with observational data, in which it tends to be inferred tautologically from the outcomes we use it to explain.[53] The power of experimental control allows us to escape this measurement challenge: we can hold dynamics of the international security environment constant and manipulate only the precise situational features of interest, while measuring dispositional characteristics using techniques borrowed from behavioral economists. Fourth, although I could use experimental methods to study the dispositional predictors of leaders' resolve, there are limits to the inferences we can draw from experiments using

51. Morgenthau 1951, 222.
52. See Bayram; Herrmann; Renshon, Lee, and Tingley; Rho and Tomz; Tingley, this issue.
53. Baldwin 1979; Sullivan 2007.

https://doi.org/10.1017/S0020818316000424 Published online by Cambridge University Press

non-elite samples to generalize to elite decision makers; restricting our focus to resolve in a public opinion context allows us to express greater confidence in the validity of the results. Uncovering evidence of individual-level microfoundations for resolve in members of the ordinary public is thus an important first step in testing a theory of resolve, and opens up the door to further investigations about leaders' resolve, a question explored further in other work.[54]

To explore individual-level microfoundations for resolve, a laboratory experiment was conducted on 317 college students recruited from undergraduate political science classes at a large midwestern research university in December 2011.[55] The study consists of three parts: a factorial experiment that manipulates the human and reputational costs of a hypothetical military intervention, a dispositional questionnaire that measures participants' time and risk preferences, and a concluding questionnaire that includes a battery of additional individual difference measures and demographic characteristics. Participants randomly received either the dispositional questionnaire or the intervention scenario first, such that the entire study can be thought of as a 2 (*ex ante* human costs: low or high) x 2 (*ex post* human costs: low or high) x 2 (reputation costs: implicit or salient) x 2 (order manipulation: scenario 1st or 2nd) fully crossed factorial experimental design. The structure of the experiment is summarized in Figure 1.

The experiment employs a foreign policy intervention scenario that has two stages. In the first (pre-invasion) stage, participants are asked whether they support a hypothetical military intervention the president wants to carry out (modified from Herrmann, Tetlock, and Visser) on behalf of a US ally under siege.[56] The scenario manipulates the amount of force the White House expects the intervention to require (low, or high), as well as reputational costs: in the reputation condition, participants are reminded "of the consequences that failing to stand up to aggression will have on America's reputation," while in the control condition, reputational concerns are not discussed. Regardless of whether participants support intervening or not, they proceed to the second stage of the scenario, in which the United States has begun the intervention. In this post-invasion stage, I manipulate the rate of casualties the United States is experiencing (low or high), along with the same reputation manipulation described earlier. In this sense, the experiment manipulates the costs of war in three different ways, reflecting three different arguments from situational theories of resolve.[57]

To explicitly capture the intertemporal quality of resolve, I employ what Gartner calls a "panel experiment": participants receive information about how the invasion

54. Kertzer 2016.

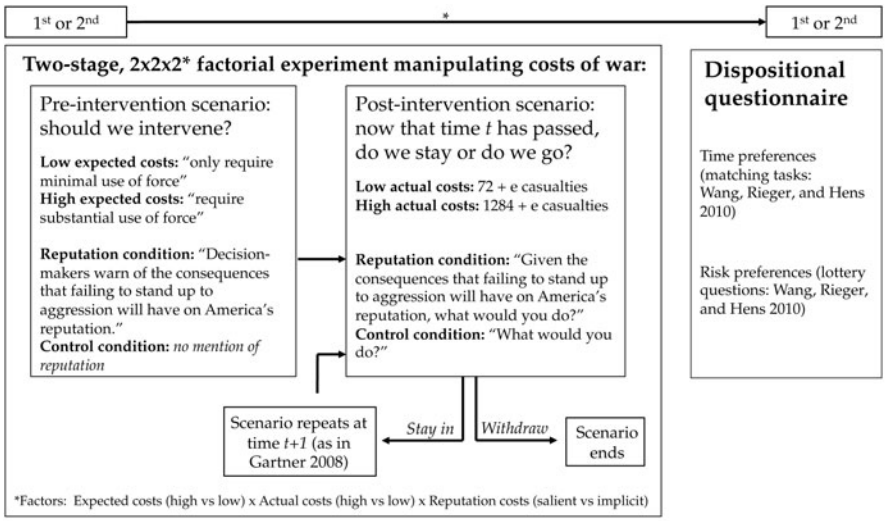
55. Participants—61 percent of whom identified as men, 77 percent as white/Caucasian, 7 percent as having served in the US Armed Forces, and 91 percent as having been born in the United States—ranged in age from seventeen to fifty-nine (μ: 21.27). For a slightly different experiment on a nationally representative sample, see Kertzer 2016.

56. Herrmann, Tetlock, and Visser 1999.

57. See Kertzer 2016.

https://doi.org/10.1017/S0020818316000424 Published online by Cambridge University Press

has progressed in the past year (the nature of which depends on the participants' experimental condition), and are asked whether the United States should withdraw its forces or not. If participants advocate withdrawal, the intervention scenario portion of the experiment ends; otherwise, the process repeats up to seven times.[58] The dependent variable is thus how much resolve participants display, as measured by the number of periods participants supported the war, analyzed using a Cox model, a semi-parametric event history model used to study survival data.[59]



Note: *Order manipulation: Intervention scenarios first and dispositional questions second, or dispositions first and scenarios second.

FIGURE 1. *Laboratory experiment structure*

To measure time preferences, I employ hypothetical choice tasks commonly used by both behavioral economists who focus on time preferences and social psychologists who study self-control.[60] I employ "matching" tasks, where participants are presented with a hypothetical choice between receiving $1,000 today, or $X at some point in the future (e.g., a year from now), and asked how much $X would have to be for it to be preferred over the immediate payment. By employing two different matching questions that vary the delay (e.g., one year versus ten years), I have enough information to construct hyperbolic discounting curves for each participant, and can calculate both the value of β (the participant's present bias) and δ (the participant's long-term discount factor), in the quasi-hyperbolic discounting function

58. Gartner 2008.
59. Box-Steffensmeier and Jones 2004.
60. For example, Ameriks et al. 2007; Wang, Rieger, and Hens 2016.

https://doi.org/10.1017/S0020818316000424 Published online by Cambridge University Press

$u(x_0, x_1 \ldots x_T) = u(x_0) + \sum_{t=1}^{T} \beta\delta^t u(x_t)$, to reflect the tendency for individuals to be more patient in the long run than in the more immediate future.

I solicit risk attitudes using a series of hypothetical lottery questions in which participants are asked how much they are willing to pay for a lottery ticket with a certain chance of receiving a particular prize.[61] The questions manipulate the value of the prize (e.g., a 60% chance of winning $100, a 60% chance of winning $400) to produce an estimate of participants' relative risk premium, a measure of risk aversion.[62] Importantly, then, the study *manipulates* situational features of military interventions, but *measures* dispositional characteristics, using survey instruments validated in other disciplines to exploit natural variation in these traits across respondents. Although this prevents me from making causal claims about the characteristics, it relieves me of the burden of attempting to concoct treatments capable of rewiring participants' dispositional makeup in precise ways.

Results

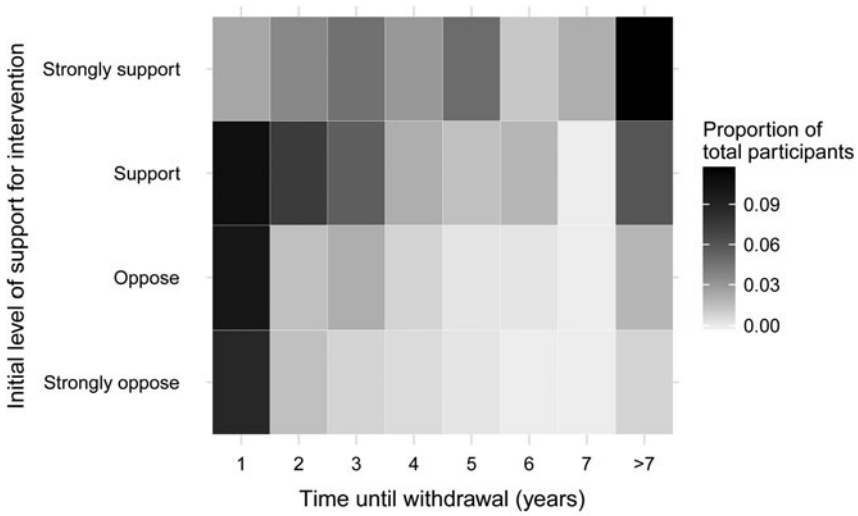
How Resolved Were the Participants?

As Figure 2 shows, there was considerable variation both in participants' initial degree of support for the intervention, and how long until they advocated for withdrawal once the US intervened. Just over 30 percent (30.4) of respondents opposed the US getting involved in the first place, compared to the 69.6 percent who supported deploying troops, but as the intervention progressed, the balance of support dropped precipitously. By the second anniversary of the intervention, nearly half (46.4%) of the respondents wanted the US to leave. Thus, respondents displayed considerable variation in resolve: around a third (32%) of participants wanted the US to cut and run by the end of its first year, whereas a fifth (20.4%) of participants continued to support the mission even when the scenario ended in its seventh year. Since, in an event history set-up, those respondents who continued to support the mission even after the scenario ended would be "right-censored" observations, participants who persisted in supporting the mission at the end of seven years were given an open-ended question asking how many more years they would continue to support the intervention. Importantly, although 37.5 percent of the responses provided specific timelines (e.g., thirteen of the sixty-four respondents mentioned ten years as a suitable benchmark for withdrawal), the rest rejected the premise of the question, giving responses like "until the job is done," or "as long as it takes."

61. Wang, Rieger, and Hens 2016.

62. Risk premia are defined as the respondent's willingness to pay for the lottery ticket, subtracted by the expected value of the lottery, divided by the expected value of the lottery.

https://doi.org/10.1017/S0020818316000424 Published online by Cambridge University Press



Note: Initial attitudes towards the mission predict the duration of support ($x^2 = 103.67$, $p < 0.000$); darker colors indicate a greater number of participants.

FIGURE 2. *Distribution of the dependent variable*

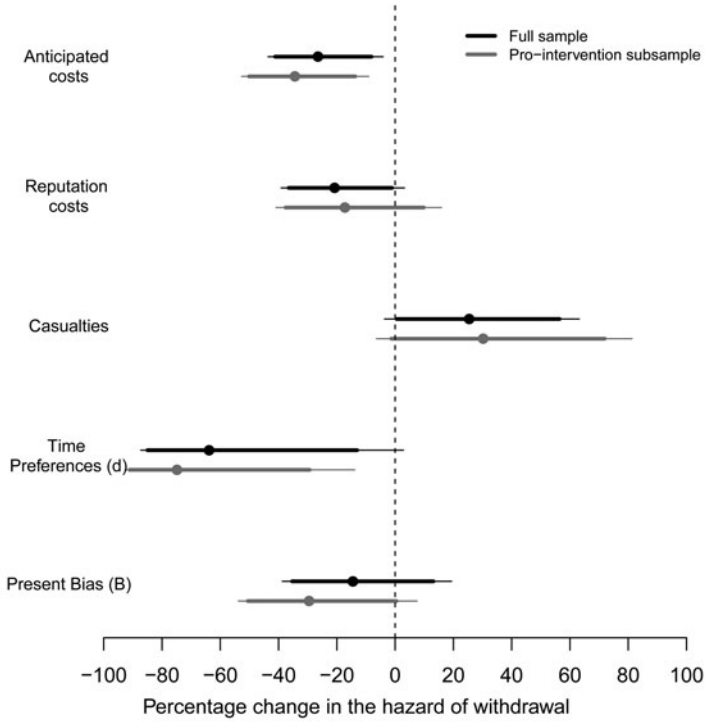
Explaining Resolve

To explain this variation, I employ semiparametric Cox models, modeling the amount of resolve participants displayed as a function of experimental treatments manipulating the costs of war, participants' dispositional characteristics, and a series of demographic controls.[63] Figure 3 depicts the results graphically, plotting a series of hazard ratios for some of the main effects of interest.

Given the two-stage nature of the experimental design—in which we have measures both of participants' desire to intervene *ex ante* and how long they support the mission *ex post*, we have the opportunity to measure resolve in a manner that closely corresponds to our second-order definition of the concept. Thus, rather than simply inferring resolve based on the length of time individuals supported the mission (deeming participants who want to immediately "cut and run" less resolute than those who want the United States to remain involved for a longer period of time), we can also look more specifically at the duration individuals who *initially* supported the mission continued to do so. After all, we know that resolve involves maintaining a policy despite temptations to back down, and the individuals who never supported the mission in the first place cannot be deemed irresolute for advocating for withdrawal. Thus, the black lines in Figure 3 present the coefficient estimates

63. Note that a survival model is necessary because of right-censoring in the data. See the online appendix for complete regression models.

https://doi.org/10.1017/S0020818316000424 Published online by Cambridge University Press



Notes: Point estimates of percentage change in the hazard ratio incurred by each of the four experimental manipulations as well as 0-1 changes in participants' δ and β parameters, with 90 and 95% confidence intervals. Negative values indicate a decreased probability of withdrawal, and thus, greater resolve.

FIGURE 3. *Coefficient plot*

for the full sample, while the gray lines depict the results for the subsample of 220 participants who initially advocated that the US intervene.

I begin by focusing on the results of the situational manipulations to establish a baseline from which to interpret the effects of time and risk preferences. As the black lines in Figure 3 suggest, casualties have a substantively large impact on participants' resolve: when casualty levels are high, participants are 31.9 percent more likely to call for troops to be withdrawn at any given period of the intervention than when casualty levels are low. Anticipated costs have the opposite effect: when participants were warned in advance that the intervention would require a substantial use of force, the respondents were 18.3 percent *less* likely to withdraw at any given point in time than if they were promised that the intervention would require only a minimal amount of force. Cost sensitivity thus appears to partially stem from actors' expectations: participants were more patient when the intervention had been framed in advance as requiring a substantial amount of force than when it required only a minimal use. Reputation costs also bolstered actors' persistence:

https://doi.org/10.1017/S0020818316000424 Published online by Cambridge University Press

participants reminded of the costs of backing down were 17.5 percent less likely to withdraw than those who received no such reminder, thereby showing that participants were attentive both to the cost of fighting and to the cost of backing down. These results are presented graphically in the dark black lines in Figure 3, which depict the hazard ratio with 90 and 95 percent confidence intervals for each of the situational manipulations; the greater the hazard of withdrawal, the less resolve participants display. Thus, the black lines in the coefficient plot show that the anticipated cost treatment and the reputation cost treatment bolster participants' resolve, while the casualty treatment lowers it.[64] Among participants who initially advocated intervening, both anticipated costs and casualties remain highly significant, with higher *ex ante* costs associated with a 26.9 percent increase in resolve, and higher casualties associated with a 34.2 percent decrease in resolve. However, the impact of reputation costs is no longer statistically significant, and is associated with only an 11.2 percent increase in resolve, although the magnitude of the change between the two models is not itself statistically significant.

Having established the situational baseline, we can then turn to the effects of our dispositional variables. Participants with more patient time preferences as measured by their long-term discounting factor (δ) displayed more resolve: an increase in δ from 0 to 1—corresponding with a change from completely discounting future costs to treating future costs as equivalent to present costs—decreases the likelihood of withdrawal by 63.9 percent. The more participants value the future, the more likely they are to persist rather than cut and run. Risk aversion displays a significant nonlinear effect: participants with a moderate degree of risk aversion ($\rho = 0.5$) are 95.1 percent more likely to withdraw than risk-neutral participants ($\rho = 0$), but past this point of risk aversion, participants actually become more resolute: a jump in risk aversion from $\rho = 0.5$ to $\rho = 1$ is associated with a 47.4 percent *decrease* in the likelihood of withdrawal, as illustrated by Figure 4.[65] Importantly, AIC scores confirm that including these dispositional variables significantly improves model fit.

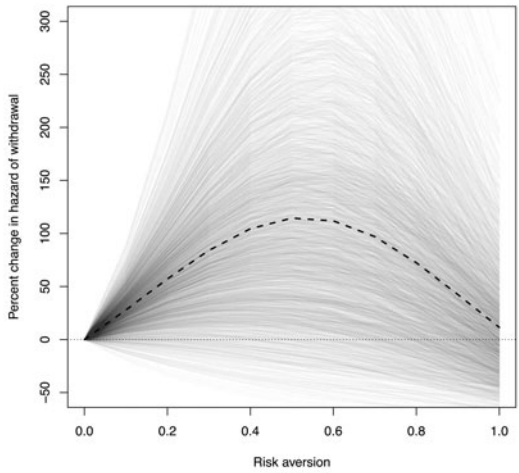
The model depicted by the gray lines in Figure 3 replicates these analyses but only for those subjects who initially advocated intervening. The effect of long-term discounting increases: a jump in δ from 0 to 1 decreases the likelihood of withdrawal at any given point in time by 74.9 percent. Similarly, β moves towards statistical significance: an increase in β from 0 to 1—which is substantively equivalent to participants becoming more patient because a quasi-hyperbolic discounting model is equivalent to an exponential discounting model when $\beta = 1$—is associated with a 29.5 percent decrease in the likelihood of withdrawal. In both cases, then, less myopic time preferences are associated with increased resolve. The impact of risk

64. The order manipulation, omitted here to save space, has no significant effect.

65. To test for nonlinearity without imposing a particular functional form, risk aversion was estimated with a penalized smoothing spline with the number of degrees of freedom determined by AIC scores; the resulting model appeared to be quadratic, and a series of likelihood ratio tests confirms that the spline fails to better fit the data than a simple quadratic ($\chi^2 = 0.518$, $p < 0.316$), and that the quadratic strongly outperforms a linear fit ($\chi^2 = 8.371$, $p < 0.004$).

https://doi.org/10.1017/S0020818316000424 Published online by Cambridge University Press

aversion is much the same as in the full sample, a curvilinear effect in which participants with a moderate degree of risk aversion ($\rho = 0.5$) are 79.3 percent more likely to withdraw than those who are risk neutral, and the extremely risk averse ($\rho = 1$) are 53.3 percent less likely to withdraw than those with moderate risk aversion.



Notes: The quadratic effect of risk aversion on participants' resolve, where each line depicts a draw from a multivariate normal distribution, given $\hat{\beta}$ and $\widehat{var}(x)$, such that darker lines indicate a more concentrated probability mass. The dashed black line represents the mean prediction.

FIGURE 4. *Quadratic effect of risk aversion*

I therefore find support for the existence of dispositional determinants of resolve: both time preferences and risk attitudes have substantively important effects on the probability of individuals choosing to "cut and run" in the military intervention, and in coherent ways. The nonlinear impact of risk aversion is notable because political scientists often collapse risk perceptions and risk preferences together, and assume that risk attitudes display a more conventional linear effect. However, just as war has many costs, it also has many risks: both withdrawing and persisting may be perceived as risky choices and, given the presence of these countervailing risk perceptions, we end up with a nonlinear relationship between risk and resolve.

Interactionist Analyses

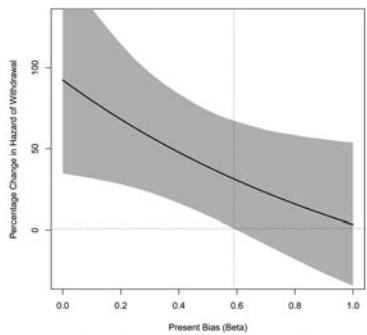
Thus far, the analyses have shown that participants' time and risk preferences significantly predict how much resolve they display. The question becomes whether, as hypothesized, these dispositional characteristics also moderate the impact of situational features.

Time preferences. I hypothesized that time preferences would moderate the impact of the costs of war on actors' resolve: since reputation costs are predominantly paid in

https://doi.org/10.1017/S0020818316000424 Published online by Cambridge University Press

the future while casualties are principally paid in the present, I hypothesized that the more patient participants' time preferences were—that is, the greater their long-term discounting (as measured by δ) and lower their present bias (as measured by β), the less sensitive they would be to casualties. Participants' δ parameters, although highly statistically significant as main effects in all of the dispositional models estimated in Tables 2 and 3 in the appendix, have no significant interaction effects with the situational manipulations. Long-term discounting and the costs of war thus appear to operate simultaneously, but in an additive rather than an interactive relationship.

In the dispositional analyses, β parameters approached statistical significance in only the pro-intervention subsample, suggesting that present bias affects resolve among only participants who wanted to intervene in the first place. However, it displays significant interaction effects with the casualty manipulation in both the full sample ($p < 0.078$) and pro-intervention subsample ($p < 0.103$), and in the direction predicted by the hypotheses: the more impatient participants are, the more the casualty manipulation makes them want to cut and run, while we fail to find evidence that relatively patient participants are significantly affected by the casualty manipulation. This interaction effect is illustrated in Figure 5, which plots the combined coefficient of the treatment effect conditional on participants' degree of present bias, with 90 percent confidence intervals, and the dashed vertical line indicating the point on the x-axis at which we can no longer distinguish the effect of the casualty treatment from 0. Notably, β did not moderate the impact of the reputation cost manipulation in either the full or the pro-intervention subsample, suggesting that time preferences moderate the impact of the costs of fighting, but not the costs of backing down.



Notes: Individuals with high present bias (low β) are less resolute (more sensitive to casualties) than more patient individuals. The dark line indicates the point estimate of the combined coefficient for the conditional effect, while the shaded gray depicts 90% confidence bands calculated using draws from a multivariate normal.

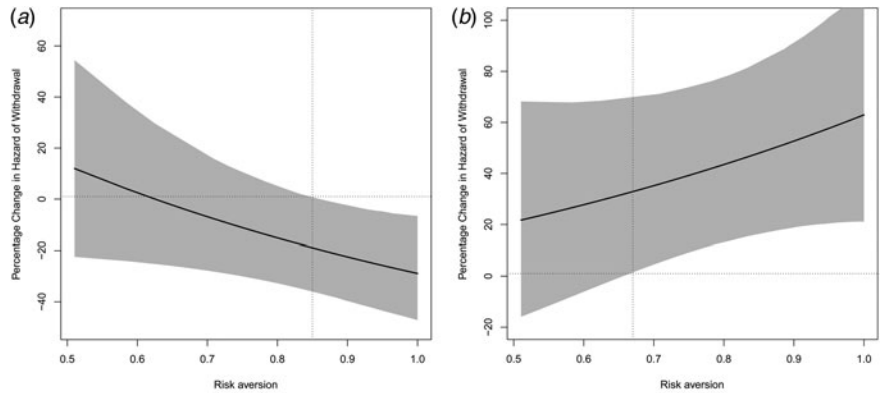
FIGURE 5. *Time preferences moderate the costs of fighting*

Risk preferences. The dispositional set of analyses showed that risk aversion was a highly significant predictor of participants' resolve, and that it displayed a quadratic effect, with resolve lowest among participants with moderate levels of risk aversion:

https://doi.org/10.1017/S0020818316000424 Published online by Cambridge University Press

as participants moved toward risk acceptance, they appeared to be more likely to want to wait things out, while as participants moved toward higher levels of risk aversion, they appeared to be more likely to want to maintain course. I suggested that this curvilinear effect stemmed from the fact that both standing ground and cutting and running can be perceived as risky, and that we would see this particular curvilinear effect if participants who were relatively risk averse were more likely to see backing down as risky than participants who were relatively risk acceptant. As Figure 6 shows, a series of interactions between risk aversion and the two main situational manipulations of interest—reputation costs and casualties—offers support for this expectation.

First, risk aversion has a statistically significant interaction with the reputation manipulation in both the full sample ($p < 0.09$) and the pro-intervention subsample ($p < 0.016$), illustrated in Figure 6(a). The more risk-averse participants are, the more persuaded they are by the reputation cost manipulation to stay the course. In contrast, we fail to reject the null hypothesis that the reputation cost treatment fails to affect resolve among those participants whose risk aversion scores are lower than $\rho = 0.85$, suggesting that as risk acceptance increases, so too does willingness to risk the reputational costs that arise from backing down in the face of aggression. Indeed, analyses on a manipulation check for perceived reputation costs find a highly significant relationship ($p < 0.01$). Risk-averse individuals are far more likely to view steep reputational costs for the US withdrawing, and the effect (+ 0.56 on a scale of 1–5) is as large as that of party ID. A parallel interaction effect is evident with respect to casualties (although only in the pro-intervention subsample: $p < 0.045$). As Figure 6(b) shows, the more risk averse participants are, the more sensitive they are to the casualty manipulation, while we fail to find that casualties significantly



Notes: Individuals who are relatively risk averse are more sensitive both to casualties and to reputation costs. The dark line indicates the point estimate of the combined coefficient for the conditional effect, while the shaded gray depicts 90% confidence bands calculated using draws from a multivariate normal.

FIGURE 6. *Risk preferences moderate both the costs of fighting, and the costs of backing down*

https://doi.org/10.1017/S0020818316000424 Published online by Cambridge University Press

lower resolve for participants whose risk aversion scores are lower than $\rho = 0.65$. Thus, for both manipulations, risk aversion's moderating effect is associated with greater caution, but in opposite directions, reflecting the extent to which the casualty treatment pushes participants out of the intervention, while the reputation treatment pulls participants in.

Conclusion

Although resolve has become one of the most fundamental concepts in international security, it has frequently played the role of what Alfred Hitchcock called a "MacGuffin": like the microfilms in *North by Northwest* or the briefcase in *Pulp Fiction*, it motivates much of the action, but its intrinsic properties are left unaddressed. I argued that we can gain better traction on resolve by lending it microfoundations, employing a novel laboratory experiment to show how time and risk preferences shape how much resolve actors display.

The experimental results provide three main sets of findings. First, although IR scholars have often treated resolve as a situational feature, throughout the results we see considerable individual-level variation: the costs of war matter—when the costs of fighting increase, or the reputational costs of backing down are rendered less salient, actors are less likely to stay the course—but our explanatory power and model fit significantly improve when we take into account characteristics of actors themselves. Actors who are relatively impatient and moderately risk averse are also more likely to be irresolute; impatient individuals are more sensitive to casualties than their patient counterparts, while risk-averse individuals are more sensitive to both casualties and reputational costs. Predicting when actors will be more or less resolved thus demands paying closer attention to actor-level attributes rather than just the environment they face.

From the perspective of IR theory, showing that dispositional characteristics matter is important because dispositional factors tend to be less directly observable than situational ones. Even if observers routinely get situational features wrong, it is nonetheless easier to know whether an actor is going to be facing high costs of war than to overcome the problem of other minds and access some interior attribute of decision makers or domestic constituencies. Thus, the greater the variation in resolve explained by actor-level characteristics, the less observable resolve is, and the less actors can rely on drawing inferences about resolve or making predictions about behavior based solely on the situations their rivals face.

If strategic actors take observable features into account when calculating and calibrating policy decisions, it follows that the more dispositional resolve is, then the more it should matter in crises because of our relative inability to take these factors into account *ex ante*. Rationalist treatments that reduce resolve to the "cost of war" therefore also reduce its potential explanatory power in the process. Relatedly, although the IR research on resolve has been largely disconnected from the growing body of research on willpower and self-control taking place outside

https://doi.org/10.1017/S0020818316000424 Published online by Cambridge University Press

political science, the major role played by the same sorts of dispositional factors that behavioral scholars routinely study when addressing similar questions in their own work suggests that resolve is not domain specific, that "political will" may be more than just a figure of speech, and that social scientists interested in resolve stand to gain by crossing disciplinary borders.

These results also raise a number of methodological issues. First, to the best of my knowledge, this is the first experimental work in IR to model both the selection into, and duration of, military operations. Although the international security literature tells us that war initiation and duration are two parts of the same continuous process, the extant experimental literature studies each step in isolation, modeling either the initial decision to use force or the question of whether troops should be withdrawn from a mission that has already begun.[66] Since the question of whether a mission is worth continuing is unlikely to be independent of the question of whether it was worth fighting in the first place, modeling both phases in one experimental design allows us to explicitly model this selection process that IR scholars frequently discuss but rarely get to directly observe.[67] Second, the fact that the dispositional sources of resolve were measured in domains unrelated to that of the intervention scenario allows for greater confidence that the results are speaking to underlying dispositions, and in the case of risk preferences, sidesteps the problem of selecting a meaningful reference point from which to judge what counts as risk averse.[68]

Third, probing microfoundations for resolve is helpful as a means of exploring the promise of the behavioral revolution in IR. Unlike approaches in IR that acknowledge heterogeneous actors only as a reluctant last resort, behavioral approaches emphasize the extent to which situational factors are indeterminate; dispositional differences matter not just because they reduce unexplained variance, but because they affect how actors perceive and define the situations they face. These findings thus echo those from others in this issue. Tingley, as well as Rathbun, Kertzer, and Paradis, show how actors in identical strategic situations respond to shifts in power differently; Bayram, as well as Herrmann, show the role that identity attachments play in structuring perceptions.[69] What is interesting about the individual differences discussed here is how amenable they are to modeling. There is nothing inherent about rationalist approaches that preclude them from placing a different methodological bet and emphasizing dispositional features because both time and risk preferences play integral supporting roles in many rational choice theories of conflict. Largely for analytic reasons, however, much of the focus has been on actors' payoff structures, so that the discipline has been able to use "the costs of war" as shorthand for an actor's resolve. My findings, then, should be seen less as challenging game theoretic models of war, and more as reminding us that dispositional features merit as much attention in making predictions as situational ones.

66. For example, Boettcher and Cobb 2006; Herrmann, Tetlock, and Visser 1999.
67. Clark 2003; Sartori 2005.
68. O'Neill 2001.
69. Bayram; Herrmann; Rathbun, Kertzer, and Paradis; Tingley; this issue.

https://doi.org/10.1017/S0020818316000424 Published online by Cambridge University Press

A number of outstanding questions remain, which similarly serve as stimulants for future behavioral work in IR. One concerns the origins of these dispositional differences. A large volume of social science research is beginning to turn to the role that formative experiences and environmental factors play in shaping individual differences like time and risk preferences, although because much of this work does not distinguish between risk preferences and risk perceptions, for example, translating its findings across domains is challenging. Another has to do with whether we see a similar degree of variation in these characteristics across elite decision makers as we do in ordinary citizens, although recent research casts considerable doubt on the idea that leaders are homogenous.[70] Given that IR scholars have (for the most part) slipped the surly bonds of the unitary actor assumption, an especially interesting puzzle concerns the question of how individual-level dispositions aggregate in groups.[71] Although an earlier tradition in economics assumed as a tenet of faith that "non standard" individual characteristics would automatically cancel each other out in groups, a later generation of work has moved beyond "invisible hand waving" and subjected the question to empirical testing.[72] Aggregation effects seem to vary with the nature of the decision-making task. In some tasks, dispositional differences average out, and group behavior is shaped by the mean level of a trait, while in others, the most extreme level of a dispositional characteristic prevails: groups tend to display a greater degree of risk acceptance than the individuals that compose them, for example, and intergroup interactions tend to be more competitive than interindividual interactions.[73] As a result of this multiplicity of mechanisms, the tendency to collapse the question into a singular form ("the aggregation problem") is likely unhelpful, obscuring more than it reveals. Unraveling these dynamics is likely to be the behavioral revolution's next frontier.

Supplementary Material

Supplementary material for this article is available at https://doi.org/10.1017/S0020818316000424

References

Ameriks, John, Andrew Caplin, John Leahy, and Tom Tyler. 2007. Measuring Self-Control Problems. *American Economic Review* 97 (3):966–72.

Axelrod, Robert. 1984. *The Evolution of Cooperation*. New York: Basic Books.

70. See, for example, Kertzer 2016; Weeks 2014.
71. See Hall and Ross 2015; and Saunders 2017.
72. Thaler 2016.
73. Moynihan and Peterson 2001; Myers and Lamm 1976.

https://doi.org/10.1017/S0020818316000424 Published online by Cambridge University Press

Baldwin, David. 1979. Power Analysis and World Politics: New Trends versus Old Tendencies. *World Politics* 31 (2):161–94.

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

Baum, Matthew A. 2002. The Constituent Foundations of the Rally-Round-the-Flag Phenomenon. *International Studies Quarterly* 46 (2):263–98.

Baumeister, Roy F., Kathleen D. Vohs, and Dianne M. Tice. 2007. The Strength Model of Self-Control. *Current Directions in Psychological Science* 16 (6):351–55.

Bénabou, Roland, and Jean Tirole. 2002. Self-Confidence and Personal Motivation. *The Quarterly Journal of Economics* 117 (3):871–915.

———. 2004. Willpower and Personal Rules. *Journal of Political Economy* 112 (4):848–86.

Benhabib, Jess, and Alberto Bisin. 2005. Modelling Internal Commitment Mechanisms and Self-Control: A Neuroeconomics Approach to Consumption-Saving Decisions. *Games and Economic Behavior* 52 (2):460–92.

Boettcher, William A., III, and Michael D. Cobb. 2006. Echoes of Vietnam? Casualty Framing and Public Perceptions of Success and Failure in Iraq. *Journal of Conflict Resolution* 50 (6):831–54.

Box-Steffensmeier, Janet M., and Bradford S. Jones. 2004. *Event History Modeling: A Guide for Social Scientists*. New York: Cambridge University Press.

Chiozza, Giacomo, and H.E. Goemans. 2004. International Conflict and the Tenure of Leaders: Is War Still Ex Post Inefficient? *American Journal of Political Science* 48 (3):604–19.

Clark, David H. 2003. Can Strategic Interaction Divert Diversionary Behavior? A Model of US Conflict Propensity. *Journal of Politics* 65 (4):1013–39.

Cline, Ray S. 1975. *World Power Assessment: A Calculus of Strategic Drift*. Boulder, CO: Westview Press.

Copeland, Dale C. 2000. *The Origins of Major War*. Ithaca, NY: Cornell University Press.

Elster, Jon. 1986. *The Multiple Self*. Cambridge, UK: Cambridge University Press.

———. 1989. *Nuts and Bolts for the Social Sciences*. Cambridge, UK: Cambridge University Press.

Fearon, James. 1994. Domestic Political Audiences and the Escalation of International Disputes. *American Political Science Review* 88 (3):577–92.

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

Fearon, James, and Alexander Wendt. 2002. Rationalism v. Constructivism: A Skeptical View. In *Handbook of International Relations*, edited by Walter Carlsnaes, Thomas Risse, and Beth A. Simmons, 52–72. London: Sage.

Frankfurt, Harry G. 1971. Freedom of the Will and the Concept of a Person. *The Journal of Philosophy* 68 (1):5–20.

Gartner, Scott Sigmund. 2008. The Multiple Effects of Casualties on Public Support for War: An Experimental Approach. *American Political Science Review* 102 (1):95–106.

Gelpi, Christopher, Peter D. Feaver, and Jason Reifler. 2009. *Paying the Human Costs of War*. Princeton, NJ: Princeton University Press.

Goemans, Hein. 2000. *War and Punishment: The Causes of War Termination and the First World War*. Princeton, NJ: Princeton University Press.

Gottfredson, Michael, and Travis Hirschi. 1990. *A General Theory of Crime*. Stanford, CA: Stanford University Press.

Grynaviski, Eric. 2014. *Constructive Illusions: Misperceiving the Origins of International Cooperation*. Ithaca, NY: Cornell University Press.

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

Hagger, Martin S., Chantelle Wood, Chris Stiff, and Nikos L.D. Chatzisarantis. 2010. Ego Depletion and the Strength Model of Self-Control: A Meta-Analysis. *Psychological Bulletin* 136 (4):495–525.

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

https://doi.org/10.1017/S0020818316000424 Published online by Cambridge University Press

Herrmann, Richard K. 1988. The Empirical Challenge of the Cognitive Revolution: A Strategy for the Drawing Inferences About Perceptions. *International Studies Quarterly* 32 (2):175–203.
———. 2017. How Attachments to the Nation Shape Beliefs About the World: A Theory of Motivated Reasoning. In The Behavioral Revolution and International Relations, supplement, *International Organization* 71 (S1):S61–S84.
Herrmann, Richard K., Philip E. Tetlock, and Penny S. Visser. 1999. Mass Public Decisions to Go to War: A Cognitive-Interactionist Framework. *American Political Science Review* 93 (3):553–73.
Herrnstein, Richard J. 1970. On the Law of Effect. *Journal of the Experimental Analysis of Behavior* 13 (2): 243–66.
Horowitz, Michael C., and Allan C. Stam. 2014. How Prior Military Experience Influences the Future Militarized Behavior of Leaders. *International Organization* 68 (3):527–59.
Jervis, Robert. 1976. *Perception and Misperception in International Politics*. Princeton, NJ: Princeton University Press.
———. 1979. Deterrence Theory Revisited. *World Politics* 31 (2):289–324.
Jost, John T., Jack Glaser, Arie W. Kruglanski, and Frank Sulloway. 2003. Political Conservatism as Motivated Social Cognition. *Psychological Bulletin* 129 (3):339–75.
Kertzer, Joshua D. 2016. *Resolve in International Politics*. Princeton, NJ: Princeton University Press.
———. Forthcoming. Microfoundations in International Relations. *Conflict Management and Peace Science*.
Kertzer, Joshua D., and Ryan Brutger. 2016. Decomposing Audience Costs: Bringing the Audience Back into Audience Cost Theory. *American Journal of Political Science* 60 (1):234–49.
Knight, Frank H. 1921. *Risk, Uncertainty and Profit*. Boston: Houghton Mifflin.
Krebs, Ronald R., and Aaron Rapport. 2012. International Relations and the Psychology of Time Horizons. *International Studies Quarterly* 56 (3):530–43.
Kydd, Andrew. 2003. Which Side Are You On? Bias, Credibility, and Mediation. *American Journal of Political Science* 47 (4):597–611.
Lyall, Jason. 2010. Do Democracies Make Inferior Counterinsurgents? Reassessing Democracy's Impact on War Outcomes and Duration. *International Organization* 64 (1):167–92.
Mack, Andrew. 1974. Why Big Nations Lose Small Wars: The Politics of Asymmetric Conflict. *World Politics* 27 (2):175–200.
Maoz, Zeev. 1983. Resolve, Capabilities, and the Outcomes of Interstate Disputes, 1816–1976. *Journal of Conflict Resolution* 27 (2):195–229.
Mattes, Michaela, and T. Clifton Morgan. 2004. When Do They Stop? Modeling the Termination of War. *Conflict Management and Peace Science* 21 (3):179–93.
McDermott, Rose. 1998. *Risk-Taking in International Politics: Prospect Theory in American Foreign Policy*. Ann Arbor: University of Michigan Press.
———. 2004. The Feeling of Rationality: The Meaning of Neuroscience for Political Science. *Perspectives on Politics* 2 (4):691–706.
Mearsheimer, John J. 2001. *The Tragedy of Great Power Politics*. New York: W.W. Norton.
Meirowitz, Adam, and Ann E. Sartori. 2008. Strategic Uncertainty as a Cause of War. *Quarterly Journal of Political Science* 3 (4):327–52.
Mercer, Jonathan. 2005. Rationality and Psychology in International Politics. *International Organization* 59 (1):77–106.
Mischel, Walter, and Joan Grusec. 1967. Waiting for Rewards and Punishments: Effects of Time and Probability on Choice. *Journal of Personality and Social Psychology* 5 (1):24–31.
Morgenthau, Hans J. 1951. *In Defense of the National Interest*. New York: Alfred A. Knopf.
———. 1985. *Politics Among Nations: The Struggle for Power and Peace*. Brief ed. Boston, MA: McGraw-Hill.
Morrow, James D. 1987. On the Theoretical Basis of a Measure of National Risk Attitudes. *International Studies Quarterly* 31 (4):423–38.

https://doi.org/10.1017/S0020818316000424 Published online by Cambridge University Press

Moynihan, Lisa M., and Randall S. Peterson. 2001. A Contingent Configuration Approach to Understanding the Role of Personality in Organizational Groups. *Research in Organizational Behavior* 23:327–78.

Mueller, John E. 1971. Trends in Popular Support for the Wars in Korea and Vietnam. *American Political Science Review* 65 (2):358–75.

———. 1980. The Search for the "Breaking Point" in Vietnam: The Statistics of a Deadly Quarrel. *International Studies Quarterly* 24 (4):497–519.

Myers, David G., and Helmut Lamm. 1976. The Group Polarization Phenomenon. *Psychological Bulletin* 83 (4):602–27.

Nelson, Stephen C., and Peter J. Katzenstein. 2014. Uncertainty, Risk, and the Financial Crisis of 2008. *International Organization* 68 (2):361–92.

O'Neill, Barry. 2001. Risk Aversion in International Relations Theory. *International Studies Quarterly* 45 (4):617–49.

Organski, A.F.K. 1968. *World Politics*. 2nd ed. New York: Alfred A. Knopf.

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

Pape, Robert A. 1996. *Bombing to Win: Air Power and Coercion in War*. Ithaca, NY: Cornell University Press.

Powell, Robert. 1987. Crisis Bargaining, Escalation, and MAD. *American Political Science Review* 81 (3): 717–35.

———. 2003. Nuclear Deterrence Theory, Nuclear Proliferation, and National Missile Defense. *International Security* 27 (4):86–118.

Ramsay, Kristopher W. 2008. Settling It on the Field: Battlefield Events and War Termination. *Journal of Conflict Resolution* 52 (6):850–79.

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

Reiter, Dan, and Allan C. Stam. 2002. *Democracies at War*. Princeton, NJ: Princeton University Press.

Renshon, Jonathan, Julia J. Lee, and Dustin Tingley. 2017. Emotions and the Micro-Foundations of Commitment Problems. In The Behavioral Revolution and International Relations, supplement, *International Organization* 71 (S1):S189–S217.

Reus-Smit, Christian. 2003. Politics and International Legal Obligation. *European Journal of International Relations* 9 (4):591–625.

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

Rosen, Stephen. 1972. War Power and the Willingness to Suffer. In *Peace, War, and Numbers*, edited by Bruce M. Russett, 167–83. Beverly Hills, CA: Sage.

Rosenau, James N. 1980. Thinking Theory Thoroughly. In *The Scientific Study of Foreign Policy*, revised ed., 19–31. London: Frances Pinter.

Sartori, Anne E. 2005. *Deterrence by Diplomacy*. Princeton, NJ: Princeton University Press.

Saunders, Elizabeth N. 2011. *Leaders at War: How Presidents Shape Military Interventions*. Ithaca, NY: Cornell University Press.

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

Schelling, Thomas C. 1966. *Arms and Influence*. New Haven, CT: Yale University Press.

Schultz, Kenneth A. 1999. Do Democratic Institutions Constrain or Inform? Contrasting Two Institutional Perspectives on Democracy and War. *International Organization* 53 (2):233–66.

Simon, Herbert A. 1985. Human Nature in Politics: The Dialogue of Psychology with Political Science. *American Political Science Review* 79 (2):293–304.

https://doi.org/10.1017/S0020818316000424 Published online by Cambridge University Press

Slantchev, Branislav L. 2003. The Power to Hurt: Costly Conflict with Completely Informed States. *American Political Science Review* 97 (1):123–33.
———. 2010. Feigning Weakness. *International Organization* 64 (3):357–88.
Slovic, Paul. 1987. Perception of Risk. *Science* 236 (4799):280–85.
Smith, Alastair, and Allan C. Stam. 2004. Bargaining and the Nature of War. *Journal of Conflict Resolution* 48 (6):783–813.
Snyder, Glenn H., and Paul Diesing. 1977. *Conflict Among Nations: Bargaining, Decision Making, and System Structure in International Crises*. Princeton, NJ: Princeton University Press.
Stam, Allan C., III. 1996. *Win, Lose, or Draw: Domestic Politics and the Crucible of War*. Ann Arbor: University of Michigan Press.
Steinbruner, John D. 1974. *The Cybernetic Theory of Decision*. Princeton, NJ: Princeton University Press.
Streich, Philip, and Jack S. Levy. 2007. Time Horizons, Discounting, and Intertemporal Choice. *Journal of Conflict Resolution* 51 (2):199–226.
Sullivan, Patricia L. 2007. War Aims and War Outcomes. *Journal of Conflict Resolution* 51 (3):496–524.
Tetlock, Philip E. 2003. Thinking the Unthinkable: Sacred Values and Taboo Cognitions. *Trends in Cognitive Sciences* 7 (7):320–24.
Thaler, Richard H. 2016. Behavioral Economics: Past, Present, and Future. *American Economic Review* 106 (7):1577–1600.
Thaler, Richard, and H.M. Shefrin. 1981. An Economic Theory of Self-Control. *Journal of Political Economy* 89 (2):392–406.
Tingley, Dustin. 2017. Rising Power on the Mind. In The Behavioral Revolution and International Relations, supplement, *International Organization* 71 (S1):S165–S188.
Toft, Monica Duffy. 2006. Issue Indivisibility and Time Horizons as Rationalist Explanations for War. *Security Studies* 15 (1):34–69.
Trager, Robert F. 2011. Multidimensional Diplomacy. *International Organization* 65 (3):469–506.
Valentino, Benjamin A., Paul K. Huth, and Sarah E. Croco. 2010. Bear Any Burden? How Democracies Minimize the Costs of War. *Journal of Politics* 72 (2):528–44.
Velleman, J. David. 2007. What Good Is a Will? In *Action in Context*, edited by Anton Leist, 193–215. Berlin: Walter de Gruyter.
Wang, Mei, Marc Oliver Rieger, and Thorsten Hens. 2016. How Time Preferences Differ: Evidence from Fifty-Three Countries. *Journal of Economic Psychology* 52 (1):115–35.
Weeks, Jessica L.P. 2014. *Dictators at War and Peace*. Ithaca, NY: Cornell University Press.
Wendt, Alexander. 1999. *Social Theory of International Politics*. Cambridge, UK: Cambridge University Press.
Wright, Quincy. 1964. *A Study of War*. Abridged ed. Chicago: University of Chicago Press.
Young, Oran R. 1968. *The Politics of Force: Bargaining During International Crises*. Princeton, NJ: Princeton University Press.

https://doi.org/10.1017/S0020818316000424 Published online by Cambridge University Press