

STRATEGIC EMPATHY AND THE SECURITY DILEMMA: Cross-National Experimental Evidence from China and the United States

Joshua D. Kertzer,¹ Ryan Brutger,² and Kai Quek³

November 11, 2018

ABSTRACT: One of the central challenges in China-US relations is the risk of a security dilemma between China and the United States, as each side carries out actions for defensively-motivated reasons, failing to realize how it is perceived by the other side. Yet how susceptible to security dilemma thinking are the Chinese and American publics? Can its deleterious effects be mitigated? We explore the individual-level microfoundations of security dilemma thinking, fielding a unique dyadic cross-national survey experiment in China and the United States. We find micro-level evidence consistent with the security dilemma, and show it is especially pronounced among Chinese respondents. We also find that IR scholars have overstated the palliative effects of empathy: perspective-taking significantly affects respondents' policy preferences, but can often lead to escalation rather than cooperation. Our findings have important implications for the study of public opinion in China-US relations, and perspective-taking in IR.

¹Paul Sack Associate Professor of Political Economy, Department of Government, Harvard University. Email: jkertzer@gov.harvard.edu. Web: <http://people.fas.harvard.edu/~jkertzer/>

²Assistant Professor, Department of Political Science, University of Pennsylvania. Email: brutger@sas.upenn.edu. Web: <http://web.sas.upenn.edu/brutger/>.

³Assistant Professor, Department of Politics and Public Administration, The University of Hong Kong. Email: quek@hku.hk. Web: <https://www.ppaweb.hku.hk/f/quek>

Peacefully managing the rise of China is the most pressing foreign policy challenge of the 21st century (Mearsheimer, 2001; Christensen, 2006). Scholars and policymakers who are worried about the risks associated with China's rise are worried about two phenomena. The first involves the hegemonic wars envisioned by power transition theory (Gilpin, 1981; Kugler and Lemke, 2000; Allison, 2017). The second involves conflicts that can occur between two security-seeking states: the notion of a security dilemma (Herz, 1950; Butterfield, 1951).

Central to the idea of a security dilemma is a dynamic in which one state carries out actions for defensively-motivated reasons, failing to appreciate how it will be perceived by the other side, leading to a spiral model of conflict that no actor actually wants (Jervis, 1978; Glaser, 1997; Booth and Wheeler, 2008; Tang, 2009; Mitzen and Schweller, 2011). Security dilemmas are a byproduct of the uncertainty generated by the anarchic structure of the international system, but they are also about failures of strategic empathy or perspective-taking — the inability of one actor to successfully put itself in the shoes of the other.

In this paper, we experimentally explore the microfoundations of perspective-taking in the security dilemma. We are interested in two questions. First of all, in an era when nationalist publics in the US and China appear set on a collision course with one another, how susceptible are the two countries to security dilemma thinking? Second, given IR scholars' persistent interest in the palliative effects of empathy in promoting international cooperation (e.g. Jervis, 1976; White, 1986; Keller and Yang, 2009; Holmes and Yarhi-Milo, 2017), what role can strategic empathy play in dampening security dilemma dynamics? If security dilemmas are about failures of empathy, does inducing perspective-taking and encouraging citizens from different countries to step into each other's shoes mitigate the security dilemma's deleterious effects? And, are naturally more empathetic people less susceptible to these misperceptions in the first place?

We present the results from a unique cross-national survey experiment fielded in China and the United States in the spring of 2016, where we present participants in both countries with a scenario in the South China Sea, manipulate the behavior of the other side, and examine its effects on policy preferences. We find micro-level evidence consistent with the logic of the security dilemma: respondents in both countries are susceptible to spiral models of conflict, viewing behavior as offensively-motivated when carried out by the other side, and defensively-motivated when carried out by their own; these effects are especially pronounced among the Chinese public. We also find that perspective-taking is a double-edged sword, because its effects are conditional on the strategic

interests at stake: experimentally inducing American participants to think about the conclusions the Chinese would draw from their behavior decreases the likelihood of endorsing escalatory policies, but doing the same for Chinese participants can actually increase the likelihood of endorsing escalation, due in part to the divergent stakes for each side. Moreover, exploiting preexisting variation in our respondents' levels of strategic empathy, we find that more strategically empathetic individuals display larger attribution asymmetries, suggesting they should be more rather than less susceptible to security dilemma thinking. In echoing other recent research questioning long-standing assumptions about the putatively liberal nature of mass public preferences (Sagan and Valentino, 2017), our findings have important implications for both theory and policy, improving our understanding of the logic of the security dilemma and the role of the mass public in either accelerating or inhibiting assertive foreign policies in Chinese-US relations, and adding a cautionary note to the voluminous body of work on the palliative effects of empathy in foreign affairs.

The discussion that follows has four parts. We begin by introducing the concept of strategic empathy, showing how in addition to the important role this cognitive type of empathy plays within strategy more generally, it plays an especially important role in psychological models of the security dilemma, a relevant challenge given growing concerns about potential security dilemmas in the South and East China Seas. We then discuss the mechanics of our cross-national experiment, before presenting our findings.

1 Strategic empathy

Strategy is about empathy — in particular, a cognitive form of empathy we define here as *strategic empathy*, or perspective-taking, defined as the ability or willingness to put oneself in others' shoes.¹ Because strategic situations are those where “the best course of action for each player depends on what other players do” (Schelling, 1960, 3, fn. 1), acting strategically requires anticipating the behavior of others. This is what sets strategies apart from plans (Freedman, 2013, xi), and

¹Shore 2014; Waldman 2014. Strategic empathy as we use it here is thus different from sympathy, which is associated with prosociality and altruism (Kertzer and Rathbun, 2015). As Head (2016, 175) notes, “empathy does not inevitably require any positive regard for the other.” This is different from how empathy is discussed in, e.g. Keohane (1984, chap. 7). For purposes of lexical variety, we use strategic empathy and perspective-taking as interchangeable with one another here, but how distinct the two are from one another varies across the psychological literature: in some work, perspective-taking is seen as a cause of empathy (Decety, 2005), while others see it as a specific subtype of empathy (Davis, 1980). We sidestep these debates for our purposes. Notably, our conception of strategic empathy and perspective-taking here is decidedly cognitive rather than affective; see Kertzer and Renshon (2014). As we note in the conclusion, future scholarship should extend our findings using affective forms of empathy instead.

game theory apart from decision theory (Morrow, 1994). Successfully predicting the actions and reactions of others requires strategic empathy or perspective-taking, the capacity to place oneself in the shoes of others. Indeed, Singer and Fehr (2005, 340) note that all of “the most fundamental solution concepts in game theory (Nash equilibrium, backward induction, and iterated elimination of dominated strategies)” depend on some form on this faculty.

In games of complete information, strategic empathy is trivial: every player already knows what the other players want and believe as a matter of design. In natural settings, however, political scientists and psychologists have pointed to a number of pertinent obstacles. Most fundamental is the “problem of other minds”: our inability to directly access the mental states of others (Hermann, 1988). In international politics, the problem is exacerbated by the anarchic structure of the international system, in which states are fundamentally uncertain about the intentions of others, who have incentives to misrepresent their private information (Fearon, 1995). Yet beyond those structural factors are a host of psychological biases that suggest that many of us operate with an “empathy deficit.”² Many of the “bounds” depicted by models of bounded rationality are bounds of empathy. Attribution biases, for example, are biases in person perception (Ross, 1977), ingroup disconfirmation biases mean we reject schema-inconsistent information about others (Larson, 1994), ethnocentric biases mean we differentially treat members of ingroups and outgroups (Kinder and Kam, 2010), and anchoring and adjustment biases mean that when most of us try to imagine how others feel in a given situation, we do so by imagining how we would feel in that situation, and then adjust outwards – usually insufficiently (Epley et al., 2004). Empathy is also an individual difference, a propensity that some of us are systematically better (or worse!) at than others (Davis, 1983; Baron-Cohen and Wheelwright, 2004).

It should not be surprising, then, that many unsavory outcomes in international politics are frequently chalked up to failures of empathy, particularly in the voluminous literature on misperceptions in international politics. Stein (1988, 249-250) notes that threat perceptions are hamstrung by a “lack of empathy”, in that “political leaders often display no sensitivity to their adversary’s sense of vulnerability while they dwell heavily on their own perception of threat.” White (1986, 62) blames Hitler invading Poland without realizing it would force Britain into the war on an “absence of empathy.” Robert McNamara offered a similar explanation for American missteps in the Vietnam War (Blight and Lang, 2004), as does Smith (2004) in regard to the unilateralism of the 2003 Iraq War.

²<http://www.npr.org/templates/story/story.php?storyId=7755013>.

Keller and Yang (2009, 181) pin the 1991 Persian Gulf War on a “failure to empathize”, manifested in the Bush administration failing to appreciate Saddam Hussein’s domestic political constraints. Holmes and Yarhi-Milo (2017, 115-116) attribute the failure of the second Camp David summit to Clinton’s failure “to build relational empathy” between Yasser Arafat and Ehud Barak.

If an absence of empathy leads to miscalculations and undesirable outcomes, it should not be surprising that IR scholars often propose empathy — in a variety of forms — as a solution to political problems. Deutschian theories of integration (e.g. Deutsch, Burrell and Kann, 1957) posit that under the right conditions, sustained interaction leads to mutual understanding and “we feeling”; security communities — regions where war is unthinkable as a means of settling disputes — are zones of institutionalized empathy (Adler and Barnett, 1998; Crawford, 2014, 550). A similar logic underlies intergroup contact theory (Pettigrew, 1998), one of the motivations behind foreign exchanges (Atkinson, 2010), and a variety of conflict resolution mechanisms (Batson and Ahmad, 2009). The constructivist tradition in IR more generally is heavily influenced by symbolic interactionism, in which social identities are constituted through a process of “perspective-taking”: “being able to anticipate other’s reactions to the Self, which requires an ability to see the Self through [others’] eyes.” (Wendt, 1999, 333).³ Finnemore and Sikkink (1998) suggest both cognitive and affective forms of empathy are common characteristics of norm entrepreneurs, while Holmes (2018) and Holmes and Yarhi-Milo (2017) argue face-to-face diplomacy is helpful because it better permits negotiators to reach mutual understanding; one of the frequently-issued arguments in favor of area studies expertise in policymaking is that a wealth of local knowledge helps decision-makers better take the other side’s perspective (White, 1986; Waldman, 2014). Measuring or manipulating empathy in a variety of ways, scholars in American political behavior have shown that more empathetic individuals are more supportive of immigration (Newman et al., 2015) and social welfare (Feldman and Steenbergen, 2001), more likely to oppose mandatory minimum sentencing (Gross, 2008), and more supportive of actions improving conditions for the Iraqi people following the 2003 Iraq War (Pagano and Huo, 2007).⁴

³Here, Wendt is careful to distinguish perspective-taking from empathy, which he defines affectively rather than cognitively, noting that “Empathy is about experiencing the Other’s feelings and welfare as if they were one’s own, about identifying with him” (Wendt, 1999, 333), which is different than our use of the construct here.

⁴See also prejudice reduction interventions, which often involve some form of perspective-taking activity - e.g. Paluck 2010; Simonovits, Kézdi and Kardos 2018.

1.1 Perspective-taking and the security dilemma

For our purposes, the potentially palliative effects of perspective-taking are particularly pertinent because of the predicament posed by a rising China. Policy-makers and political scientists preoccupied with parsing China's rise (e.g. [Christensen, 1999](#); [Mearsheimer, 2001](#); [Johnston, 2003](#); [Medeiros and Frelv, 2003](#); [Friedberg, 2005](#); [Gries, 2005](#); [Goldstein, 2005](#); [Christensen, 2006](#); [Ikenberry, 2008](#); [Ross, 2009](#); [Schweller and Pu, 2011](#); [Chen, Pu and Johnston, 2014](#); [Johnston, 2017](#)) typically focus on one of two phenomena.⁵

The first is the specter of hegemonic war. As [Gilpin \(1981\)](#) argued, the international political order tends to reflect the interest of the most powerful states in the system; as the balance of power shifts due to the law of uneven growth, the newly powerful will attempt to change the system to better reflect their interests. Declining hegemony, then, have an incentive to wage preventive war in order to forestall the rising power's rise ([Levy, 1987](#)), such that periods of power transition have historically tended to be periods of instability ([Kugler and Lemke, 2000](#); [Allison, 2017](#)). Importantly, however, the dire predictions of power transition theory are contingent upon the rising power being a revisionist state, seeking to overturn the established order. If China is invested in the existing international political order — an arrangement that enabled it to lift nearly three quarters of a billion people out of poverty over the past several decades — then no conflict should occur.

The second is the security dilemma ([Herz, 1950](#); [Butterfield, 1951](#)). As [Booth and Wheeler \(2008, 4-8\)](#) and [Tang \(2009\)](#) note, IR scholars use the term to refer to a wide range of phenomena, but for our purposes, it begins with the premise that because of the anarchic structure of the international system, states are fundamentally uncertain about the intentions of others ([Waltz, 1979](#); [Copeland, 2000](#); [Mearsheimer, 2001](#)). When one state takes actions that are intended only to bolster its own security, other states are likely to assume the worst, and erroneously interpret a defensively-motivated action as an offensively-motivated one. The result is what [Jervis \(1976\)](#) calls the “spiral model” of conflict, where war no one wants emerges as a result of misperceptions that neither side can shake off ([Glaser, 1997](#); [Tang, 2009](#); [Mitzen and Schweller, 2011](#)). Thus, whereas hegemonic war only occurs when one of the actors is a revisionist state, the security dilemma can occur even among two states supportive of the status quo. Concerns about potential security dilemmas in East Asia — whether

⁵The 2012 TRIP survey suggests the strategic importance of peacefully managing China's rise is shared by both academics and policymakers alike, topping the list of foreign-policy problems facing the United States provided by both IR scholars in the United States, and national security practitioners within the US government. See https://www.wm.edu/offices/itpir/_documents/trip/ivory_tower_survey_2012.pdf.

between China and the United States, or between China and America's allies in the region, Japan and South Korea — thus loom large in much of the East Asian security scholarship (e.g. [Christensen, 1999, 2002](#); [Liff and Ikenberry, 2014](#)). American scholars have noted that “China and the US seem to be locked in a complex security dilemma with no easy solutions in sight” ([Teng, 2017](#)) with the ambiguity of President Trump's rhetoric about the region leading to a “growing security dilemma which could spiral into a regional arms race, destabilizing Asia and increasing the chance of conflict if there is not a swift shift in direction” ([Ludwig, 2017](#)). Yet Chinese scholars have recently made similar arguments, suggesting that “the South China Sea dispute has become the “security dilemma” of the two countries” (南海成为两国的“安全困境”), or looking for ways for the US and China to “step over the trap of “security dilemmas” in the South China Sea” (并进而跨越南海“安全困境”的陷阱).⁶

What is interesting about both of these worst-case scenarios is that they carve out an important role for strategic empathy, or perspective-taking. Because hegemonic wars depend on the rising power harboring revisionist intentions, debates in Washington over how the United States should handle a rising China — whether we should engage China and enmesh it in international institutions ([Ikenberry, 2008](#)), reassure China by using reciprocal gestures to reduce tensions ([Steinberg and O'Hanlon, 2014](#)), contain it militarily ([Mearsheimer, 2014](#)), or some combination thereof — are partially debates about future states of the world (e.g. how much will Chinese economic growth slow? How much will Chinese nationalism rise?), but are also debates about what China wants ([Johnston, 2003](#)). At their heart, then, American grand strategy debates over the Asia-Pacific hinge on the fundamental challenge of overcoming “the problem of other minds” and accurately assessing the motivations of others ([Morgenthau, 1985](#); [Herrmann, 1988](#)). As [Butterfield \(1951, 154\)](#) noted when coining what he referred to as the “irreducible dilemma”, the same attributional challenge serves as the core of the security dilemma. In his influential discussions of the security dilemma, [Jervis \(1978, 181\)](#) discusses the environmental features that makes security dilemmas more likely to occur, but also explicitly attributes the dilemma to “failures of empathy”, and notes that “empathy and skillful statesmanship can reduce this danger” (p. 212).⁷ An ability to put oneself in others'

⁶滕建群, 滕建群: 美国想赢有心无力 (“Teng Jianqun: America desires to but is unable to win”), *China Institute of International Studies*, 2018. http://www.ciis.org.cn/chinese/2018-01/17/content_40190420.htm. 吴士存, 中美亟待破解南海“安全困境” (“An Imperative for China and the U.S. to Solve the “Security Dilemma” in South China Sea”), *National Institute for South China Sea Studies*. 2016. http://www.nanhai.org.cn/review_c/175.html.

⁷Similarly, [Glaser \(1997, 174\)](#) notes that the magnitude of the security dilemma is reduced if “states can rely on unit-level information about others' motives”; signals of reassurance can provide this information, but so too does empathic accuracy, particularly if costly signals do not speak for themselves.

shoes is also central to what [Booth and Wheeler \(2008\)](#) call “security dilemma sensibility”, crucial to breaking out of the dilemma.

Some of these challenges explicitly implicate strategic empathy at the level of political elites. For example, if the theoretical literature referenced above is accurate, maintaining a robust diplomatic corps with deep area-specific knowledge should enable decision-makers to more accurately understand how the Chinese government perceives their country’s national interests, helping American political leaders make better decisions in both long-term planning and short-term crises ([White, 1986](#)). Others, though, also implicate the mass public, the population we study here. This is obviously the case in democratic countries like the United States, where the linkages between public opinion and foreign policy are more solidified than in non-democratic regimes like China ([Holsti, 2004](#)). As [Kertzer \(2016, 50-51\)](#) argues, public opinion shapes foreign policy in democracies through four pathways. IR scholars typically focus on the direct pathway (what [Aldrich et al., 2006](#) call “the electoral connection”), in which voters punish incumbents for foreign policy misadventures ([Fearon, 1994](#)), but public opinion also exerts influence on foreign policy through indirect pathways. One is “latent” public opinion, in which strategic politicians take expected public reactions into account ex ante when formulating their policy choices ([Saunders, 2015](#)); this is one reason why troop contributions to foreign military interventions tend to oscillate with election cycles ([Marinov, Nomikos and Robbins, 2015](#)). Another works through domestic opposition groups, who are more likely to capsize the president’s broader legislative agenda when her foreign policy is unpopular ([Gelpi and Grieco, 2015](#)). Another works through foreign decision-makers, who know threats and promises coming out of the White House are less credible when the mass public is not behind the President ([Kertzer, 2016, 51](#)). In the specific case of US-China relations, Pew has found that negative attitudes towards China have increased by 26 percentage points in the United States since 2006,⁸ and President Trump made lambasting China a central part of his campaign speeches in the 2016 Presidential election. In experimental research, [Tingley \(2017\)](#) finds that many members of the American public tend to think of a rising China using the same commitment problem framing that underlies models of preventive war. Understanding how susceptible the US public is to security dilemma thinking is a similarly worthy endeavor.

Yet understanding whether similar patterns exist in Chinese public opinion is also important. One major concern among Western pundits is that the Chinese Communist Party’s domestic legit-

⁸<http://www.pewresearch.org/fact-tank/2017/02/10/americans-have-grown-more-negative-toward-china-over-past-decade/>

imacy is tied to maintaining high levels of economic growth; as growth rates inevitably slow, the regime may instead fan the flames of nationalism in the mass public, redirecting discontent away from Beijing and towards foreign actors like the United States (Friedberg, 2005, 29-30). It is for this reason that outside observers have been so fixated on nationalist sentiment in the Chinese public (Weiss, 2014; Johnston, 2017), particularly if the Chinese government cultivates nationalist protests at home as a way to tie its hands in international negotiations abroad (Weiss, 2013, see also Gries, 2001; Wallace and Weiss, 2015). Ross (2009), for example, argues that China's maritime buildup is attributable to "naval nationalism" in a public that associates great power status with maritime grandeur. More generally, a recent strand of research on authoritarian accountability suggests that the Chinese government is sensitive to public sentiment in a variety of ways (e.g. Chen, Pan and Xu, 2016; Distelhorst and Hou, 2017; Truex, 2016; Meng, Pan and Ping, 2017). Indeed, despite maintaining an authoritarian regime, the Chinese government spends considerable effort monitoring public opinion; Weiss and Dafoe (2016) note that the Chinese government employs two million analysts to monitor public sentiment on Chinese social media, precisely because it is aware of the outsized role that domestic unrest has historically played in response to international disputes. Better understanding the dynamics of Chinese public opinion towards the use of force (e.g. Reilly, 2011; Li, Wang and Chen, 2016; Weiss and Dafoe, 2016; Bell and Quek, 2018) is thus a politically consequential task.

1.2 The dark side of strategic empathy

Does strategic empathy have the potential to dampen distrust and mitigate security dilemma thinking between the United States and China? Much of the literature in both IR and American politics summarized above, by pointing to the positive effects of perspective-taking on building trust and reducing misperceptions, would answer in the affirmative. If the security dilemma is indeed a "failure of empathy", redressing this surfeit should have beneficial effects. According to this conventional wisdom, which we dub the *palliative effects of empathy* hypothesis, more strategically empathetic individuals should be less prone to misperceptions, and more likely to take other perspectives into account. Encouraging individuals to think about the situation from the perspective of the other side should lead them to exercise more caution and prudence.

Yet there are reasons to be skeptical about seeing strategic empathy solely through liberal lenses. First, if empathy is an inherent part of strategy, it is because considering the potential reactions of

others is part of defining the strategic environment one faces. It does not, however, transform the underlying nature of that environment; it can illuminate the existence of complementary interests where they exist, but cannot transform conflicting interests into harmony (Axelrod and Keohane, 1985). In an innovative set of experiments, Caruso, Epley and Bazerman (2006) (see also Epley, Caruso and Bazerman, 2006) studied perspective-taking in the context of two classic games in the social sciences: the prisoners' dilemma, and trust games. In the prisoners' dilemma experiment, participants in the control condition were administered the classic prisoners' dilemma game, where participants are shown a payoff matrix and asked whether they would like to cooperate or defect (the dominant strategy, but which leaves players collectively worse off). Participants in the treatment condition completed the same task, but were also given a perspective-taking manipulation that encouraged them to first consider the thoughts of their opponent; participants who were encouraged to engage in perspective-taking were twice as likely to defect as those in the control. In the trust game, participants in the control condition were administered the standard trust game from behavioral economics, in which participants are given a sum of money, and have to choose how much to allocate to another player, who invests the money and then decides how much of the proceeds to return to the original player; the initial amount allocated is used as a measure of trust. When a perspective-taking treatment was added, players were significantly *less* trusting than in the control. In this sense, perspective-taking can make cooperation *less* likely, by heightening our awareness of conflicts of interest. Pierce et al. (2013, 1986) similarly report that in competitive contexts, "perspective taking is akin to pouring gasoline on a fire", effectively transforming "the Golden Rule from "do unto others as you would have them do unto you" to "do unto others as you think they will try to do unto you.""

Second, perspective-taking does not actually overcome "the problem of other minds"; a propensity or willingness to engage in empathy does not mean the inferences drawn will be accurate. As (Jervis, 1976, 112-113) notes, given the propensity for individuals to engage in "mirror-imaging" (Bronfenbrenner, 1961), a mere willingness to empathize with an adversary is insufficient. Actors:

"... must try to empathize with a variety of possible outlooks, any one of which could be a true representation of the adversary. It is not enough to calculate how the other will respond to your action if your image of him is correct. You must also try to estimate how the other will respond if he has intentions and perceptions that are different from those that you think he probably has. (In doing so you must also keep in mind that he

is likely to think that you do understand him and so will view your policy as though it were designed to deal with his own policy as he, rather than you, sees it.)"

Epley and Caruso (2008) point to a series of mechanisms to explain why empathic accuracy is so difficult. First are *failures of activation*: individuals must be willing to think about the perspective of other actors. Second are *miscalibrated* and *inaccurate adjustment*: even if individuals are willing to think about the perspective of another actor, they must then find a way to decouple this assessment from their own perspective, which requires making use of additional information, which may not necessarily be accurate. As a result, perspective-taking in practice is a form of egocentric anchoring and adjustment, where individuals engage in perspective-taking by anchoring on their own perspective — imagining how others would act by first imagining how they themselves would act — and then trying to “get over themselves” by adjusting that assessment outwards, usually insufficiently (Epley and Gilovich, 2006; see also Kertzer and Renshon, 2014). Moreover, as the discussion of area studies above suggests, taking another actor’s perspective requires utilizing knowledge about the actor, and we always know less about others than we do about ourselves. Drawing mental inferences about dissimilar or unfamiliar targets can therefore induce us to rely on stereotypes (Ames, 2004). If the stereotypes we have about other actors are negative ones — as is often the case in conflictual situations in international politics — the act of engaging in perspective-taking need not necessarily lead to positive outcomes. Sassenrath, Hodges and Pfattheicher (2016) argue that taking the perspective of a threatening target can lead to more negative evaluations; Tarrant, Calitri and Weston (2012) show this effect is especially likely among high ingroup identifiers. In contrast with the palliative effects hypothesis, then, we might imagine a *contingent effects* hypothesis, where the effects of strategic empathy depend upon the strategic setting. If strategic empathy affects behavior by increasing actors’ attentiveness to the situational features they face, acting more strategically need not imply acting more cooperatively. In situations where conflicts of interest exist, or actors have negative stereotypes about one another — in other words, the very conditions where IR scholars usually claim strategic empathy is most needed — perspective-taking can exacerbate competition rather than induce cooperation.

2 Methods

To explore the potential effects of strategic empathy in the context of US-China relations, we conducted a pair of cross-national survey experiments in China and the US in the spring of 2016, in which we presented an identical experimental scenario to national samples in both countries at the same time. In the United States, we fielded our experiment in English, embedding it in a survey on a national sample of 1822 American adults recruited by Survey Sampling International (SSI), stratified based on census targets for gender, age, household income, and education.⁹ In China, we fielded our experiment in Chinese, embedding it in a national survey that was fielded online to cover all provinces and capital municipalities in Mainland China. Anonymized online surveys are known to reduce social desirability biases and improve response validity, which is particularly important in sensitive environments such as China (e.g. [Chang and Krosnick, 2010](#); [Huang, 2015](#), 630).¹⁰ To conduct our experiment in China, we partnered with a survey company to recruit a sample of 1,556 Chinese adults (18 years and older) matching the 2010 National Census on gender, age, race, income and geography. Because these subjects were directed to the Qualtrics survey platform at the researcher end, we maintained full control over our experiment and the data collection.

Respondents in both countries were presented with a parallel experimental scenario, involving tensions in the South China Sea between the United States and China.¹¹ In each experiment, we manipulated the behavior of the other side: thus, respondents in the United States were presented with actions by the Chinese, while respondents in China were presented with actions by the United States. This parallel dyadic structure is one of the innovations of our methodological approach. Although there have been a number of important survey experiments in IR fielded in multiple countries (e.g. [Tomz and Weeks, 2013](#)), they tend to be monadically rather than dyadically structured. Indeed, it is striking that whereas the original audience cost model is dyadic ([Fearon, 1994](#)), most of the public opinion research testing its microfoundations, like most survey experiment work more generally, has been monadic (e.g. [Tomz, 2007](#); [Trager and Vavreck, 2011](#); [Kertzer and Brutger, 2016](#)). In this sense, one of our efforts here is to bridge the divide between the crisis bargaining literature, and experimental studies of public opinion, and to follow the exhortations of [Johnston \(2012\)](#), [Bell and Quek \(2018\)](#) and others in directly testing our theories of IR in an East Asian

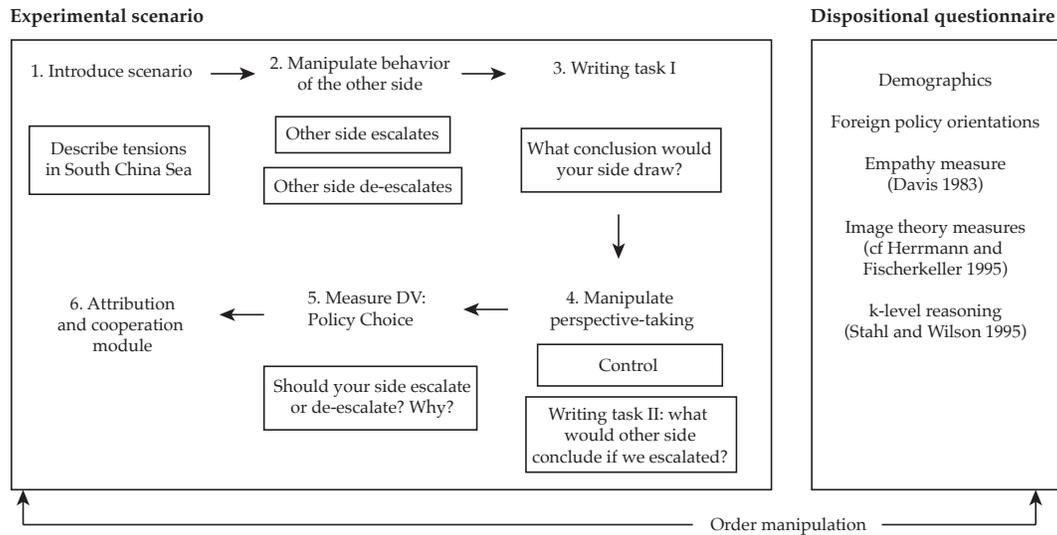
⁹For other recent experiments in political science fielded on SSI, see [Brutger and Kertzer \(2018\)](#); [Quek \(2017\)](#).

¹⁰For more on how our study addressed social desirability bias, see our discussion in Appendix §1.

¹¹For more details on the survey instrument and translation procedures, see Appendix §2.

context.

Figure 1: Study design



The overall structure of the experiment is presented in Figure 1, and the Chinese and English text used in our experiments are reproduced alongside each other in Appendix §3. Respondents in both countries began the experimental scenario by reading: “Recently there has been much attention over tensions in the South China Sea. Multiple countries in the region have claimed rights to disputed international waters, which are home to a wealth of natural resources, fisheries, and trade routes — all of which are at stake in the increasingly frequent diplomatic standoffs. [China/The United States] is concerned about the [United States’/China’s] assertiveness in the region.” That is to say, respondents in China were told about the United States, while respondents in the United States were told about China. In this sense, the experimental design is both parallel and asymmetric: although the structure of the experiment is identical across both samples, the stakes are not, since as its name suggests, the South China Sea is closer to China than to the United States, and is of significantly greater salience for the former than the latter. The research design thus deliberately captures one of the key challenges underlying US security concerns in East Asia, in which many of the potential disputes the US worries about take place in another great power’s backyard; this is the strategic context commentators focus on when discussing potential security dilemmas in US-Chinese relations. In this sense, the experimental design has greater internal validity and mundane realism than one about a hypothetical dispute equidistant from each country’s mainland.

Respondents were randomly presented with one of two different actions carried out by the other side. Half the sample considered a scenario in which the the other side decided to *decrease* naval deployments in strategic maritime zones in the South China Sea, whereas the other half of the sample considered a scenario in which the other side decided to *increase* deployments. To strengthen the treatment, participants were also given a writing task asking them to write out what conclusions their country would draw from that action.

Participants in the control condition were then presented with our main dependent variable, a measure of respondents' policy preferences in the scenario: the extent to which they wanted their country's military activities to increase or decrease, administered using a branching item used to produce a seven-point scale ranging from 1 (decrease a lot) to 7 (increase a lot).

Prior to being administered the dependent variable, participants in the perspective-taking treatment condition, however, were induced to engage in perspective-taking. Participants in the treatment condition were given an additional writing task, asking them to write what conclusions *the other side would draw* if their own side were to escalate: that is, Chinese participants in the treatment condition were asked what conclusions the Americans would draw if China were to escalate, while American participants in the treatment condition were asked what conclusions the Chinese would draw if the United States were to escalate. Participants in the treatment condition were then administered the same dependent variable administered to participants in the control. Participants in both conditions were also given an open-ended response asking them to explain their answer.¹²

Finally, participants in the experiment were administered a follow-up module where they attributed motivations for the actions of each country if each were to increase its presence in the South China Sea (each of which are administered on a seven point scale ranging from 1 ("for purely defensive reasons") to 7 ("for purely for offensive reasons")).

In addition to the experimental scenario, respondents also completed a dispositional questionnaire alongside the usual battery of demographic characteristics. Most notably, the dispositional questionnaire included a dispositional measure of strategic empathy: the perspective-taking subscale of Davis's (1983) interpersonal reactivity index (IRI), a commonly-used cognitive measure of empathy in the psychological literature. Whereas the perspective-taking treatment experimentally

¹²The study thus can be thought of as an encouragement design, in that participants were experimentally encouraged to engage in perspective-taking, but some individuals may have already been engaging in perspective-taking in the control, while others who were administered the treatment may fail to comply with it. We consider the implications for the interpretation of our findings in Appendix §4.1, showing it renders the results reported in the main text a more conservative test.

encourages half of our participants to engage in perspective-taking, the perspective-taking scale measures perspective-taking as a trait, scoring participants based on the extent to which they agree with questions like “I believe that there are two sides to every question and try to look at them both.” In this way, we can test both whether making participants engage in perspective-taking changes their policy preferences, but also whether participants who are naturally more empathetic advocate for different policy choices than those who are naturally less empathetic. To avoid downstream effects, these dispositional questions were randomly assigned in an order manipulation, such that some participants answered the questionnaire prior to the scenario, and others after.

Although the study design is somewhat elaborate, it enables us to study four theoretically valuable quantities of interest:¹³

1. By comparing the average policy choices advocated by American and Chinese respondents in each experimental cell, we can calculate the baseline level of escalatory preferences advocated by citizens in each country.
2. By estimating the treatment effect of perspective-taking on policy choices, we can test whether experimentally inducing participants to take the perspective of the other side causes them to favor more or less escalatory policies in each country.
3. By estimating the treatment effect of perspective-taking on the explanations respondents give for their policy choices, we can test whether experimentally inducing participants to take the perspective of the other side changes the kinds of justifications they offer for their policy preferences.
4. By comparing the attributions respondents offer for Chinese and American escalation in the attribution module, we can measure propensity for security dilemma thinking by calculating an “attribution asymmetry” measure at both the sample- and respondent-level, indicating how much more offensively-motivated citizens perceive an action to be when it is carried out by the other country than by their own.

¹³It also permits us to study a fifth quantity of interest — whether predispositional levels of empathy are associated with favoring more or less escalatory policies in each country. Due to space constraints, we present these results in Appendix §4.3.

3 Results

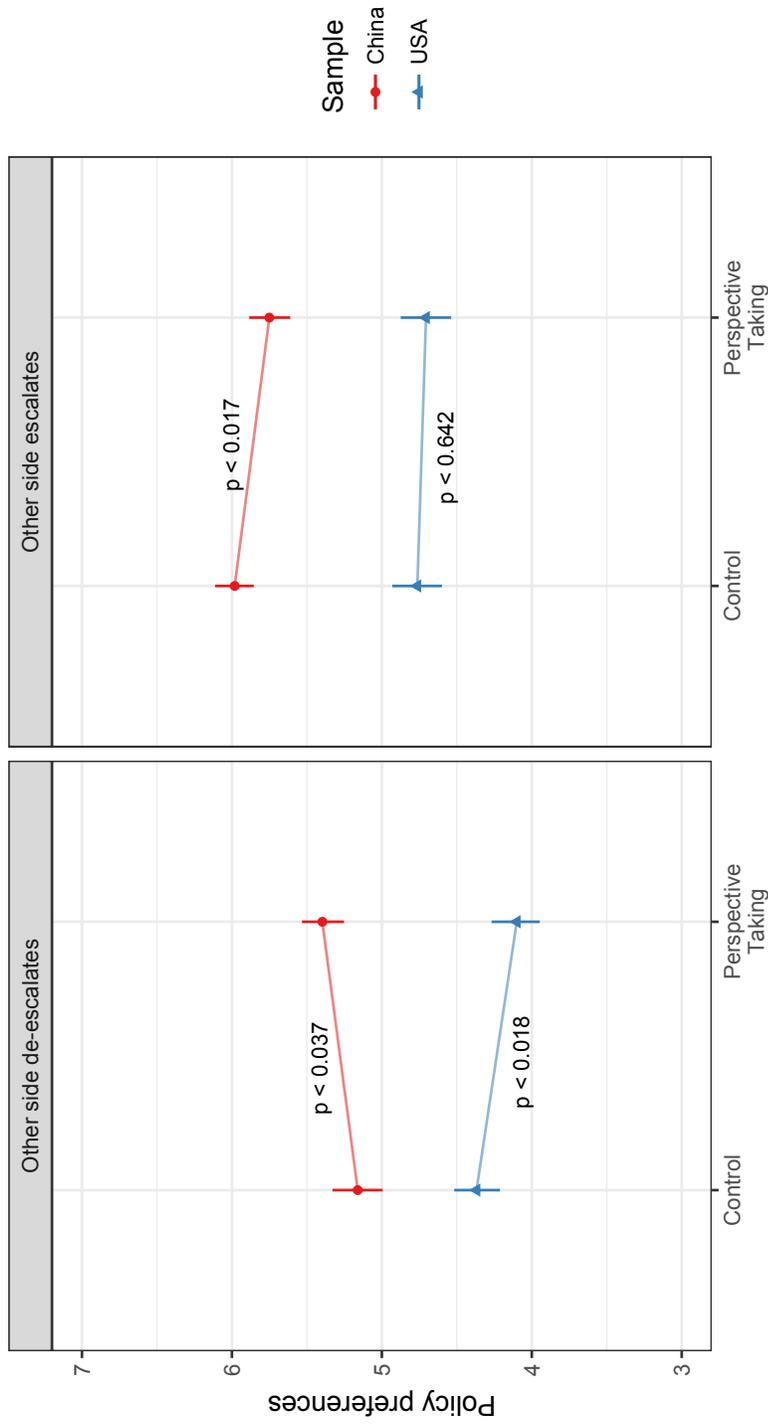
We structure our presentation of the results around the four quantities of interest enumerated above. We begin by presenting the cell means and estimating the average treatment effects of the perspective taking treatment on respondents' policy preferences in each country, showing how perspective-taking encourages reciprocity among American respondents, but opportunistic behavior among Chinese respondents; we obtain similar findings using a series of Structural Topic Models (STMs) to explore how the perspective-taking treatment affects the explanations participants offer for their policy choices. We show that this asymmetric pattern is due to strategic empathy's effects being contingent upon the contours of the strategic setting, rather than to compositional or relational differences between the two samples. Finally, we look at the question of security dilemma thinking from another angle, using the attribution questionnaire to calculate sample- and respondent-level measures of attribution asymmetry, showing that respondents tend to see actions as significantly more offensively-motivated when carried out by the other side, especially in China, and that these attribution asymmetries are strongest amongst our more empathetic respondents.

3.1 Cell means and average treatment effects

To examine the the relationship between empathy and the security dilemma, our vignette focuses respondents' attention on the South China Sea and the security policies of the US and China. Our primary dependent variable is respondents' preferred policy with regard to military escalation in the South China Sea. Higher values indicate respondents wanted their government to increase its military activities, while lower values indicate a preference for decreasing military activities. Since the dependent variable follows both the escalation (de-escalation) treatment and the perspective taking treatment in our experiment, Figure 2 presents the average policy preferences for respondents in each country, conditioning on treatment assignment.

As Figure 2 shows, the baseline preference of military escalation significantly differs between the Chinese and American publics. Chinese respondents are significantly more likely to support increasing their military activities in the South China Sea, with their average preferred policy scoring 1.07 ($p < 0.001$) points higher than the US respondents on a seven point scale. Regardless of treatment condition, the Chinese public prefers substantially higher levels of escalation in the South China Sea than their American counterparts, which is consistent with the conflict's proximity and

Figure 2: Effects of treatments on policy preferences



Note: Higher values indicate more escalatory policy preferences. The figure shows that respondents in both countries are more supportive of escalation when the other side does the same, but that the Chinese public are consistently more supportive of escalation than the Americans are. Interestingly, the effect of the perspective-taking (depicted here with p-values from t-tests) is highly contingent. Consistent with palliative accounts of perspective-taking in IR, perspective taking makes Americans less escalatory when China does the same, facilitating positive reciprocity. However, for Chinese respondents the effect of perspective-taking is inverted: when the United States de-escalates, the Chinese respond to the perspective-taking treatment by favoring greater escalation, while when the United States escalates, the Chinese respond to the perspective-taking treatment by favoring greater de-escalation.

stakes for the Chinese, a point we return to below.

We will only briefly mention the effects of our escalation treatment, since it is of less theoretical interest. By comparing the left and right panels of Figure 2, we find that the public in both countries responds in a manner consistent with the spiral model of escalation that fuels the security dilemma. In both samples, respondents prefer higher levels of escalation by their own government after the other country escalates their military activities, with the average increase for the US and China being 0.50 and 0.59 respectively (both significant at $p < 0.001$). This pattern of treatment effects is not surprising, but shows that respondents are reacting in a coherent manner to the escalation treatment, consistent with pre-existing theories of tit-for-tat escalation.

More interesting are the effects of our perspective taking treatment on policy preferences, which has divergent effects in the US and China. In the US, the perspective taking treatment causes participants to advocate for de-escalation when the other side is observed decreasing their military activities ($-0.27, p < 0.018$).¹⁴ However, we find that the perspective taking treatment has the opposite effect among the Chinese public. After engaging in perspective taking and observing the US de-escalate its military activities, Chinese respondents prefer significantly higher levels of military activities ($0.23, p < 0.037$). This finding illustrates that increased empathy is not a panacea for alleviating the security dilemma, and instead our perspective taking treatment encourages Chinese respondents to opportunistically seek to take advantage of their adversary’s de-escalation by moving to gain a decisive advantage in the conflict. Interestingly, the perspective taking treatment also has the effect of decreasing the hawkishness of the Chinese public’s preferred policy when the US has already escalated its military activities ($-0.23, p < 0.017$), suggesting that it makes the Chinese public more likely to be deterred from military escalation when the other side escalates, somewhat mitigating the preference for mutual tit-for-tat escalation.

3.1.1 Effects of perspective taking on policy reasoning

The above analysis shows that perspective-taking does not necessarily have palliative effects: perspective-taking induces reciprocity among American respondents – who are more likely to de-escalate if the Chinese do the same – but also induces divergent behavior among our Chinese respondents, who seek to project strength when the US de-escalates, but de-escalate in the face of American strength. In

¹⁴Given potential non-compliance, in Appendix §4.1 we combine sensitivity analyses with an instrumental variable approach to estimate causal average complier effects, showing how the results presented in the main text are likely underestimating the effect of the treatment.

Appendix §4.3, we offer further evidence against the palliative effects of empathy hypothesis, showing that when we exploit natural variation in respondents' levels of strategic empathy, more empathetic respondents are never significantly less escalatory (and in the Chinese case, are sometimes more escalatory).

A similar interpretation can be gleaned from the explanations respondents gave for their preferred policy choice. To do this, we asked all respondents to explain their preferred policy choice in a free response format. We analyze these open-ended responses using a structural topic model (STM), a semiautomated text analysis model that allows us to measure the impact of our perspective taking treatment on the language used by respondents. Thus, whereas the previous analysis asks how experimentally inducing perspective-taking affects *what* respondents want their countries to do, this analysis turns to discourse, asks how perspective-taking affects *why* respondents want their countries to behave in a particular way.

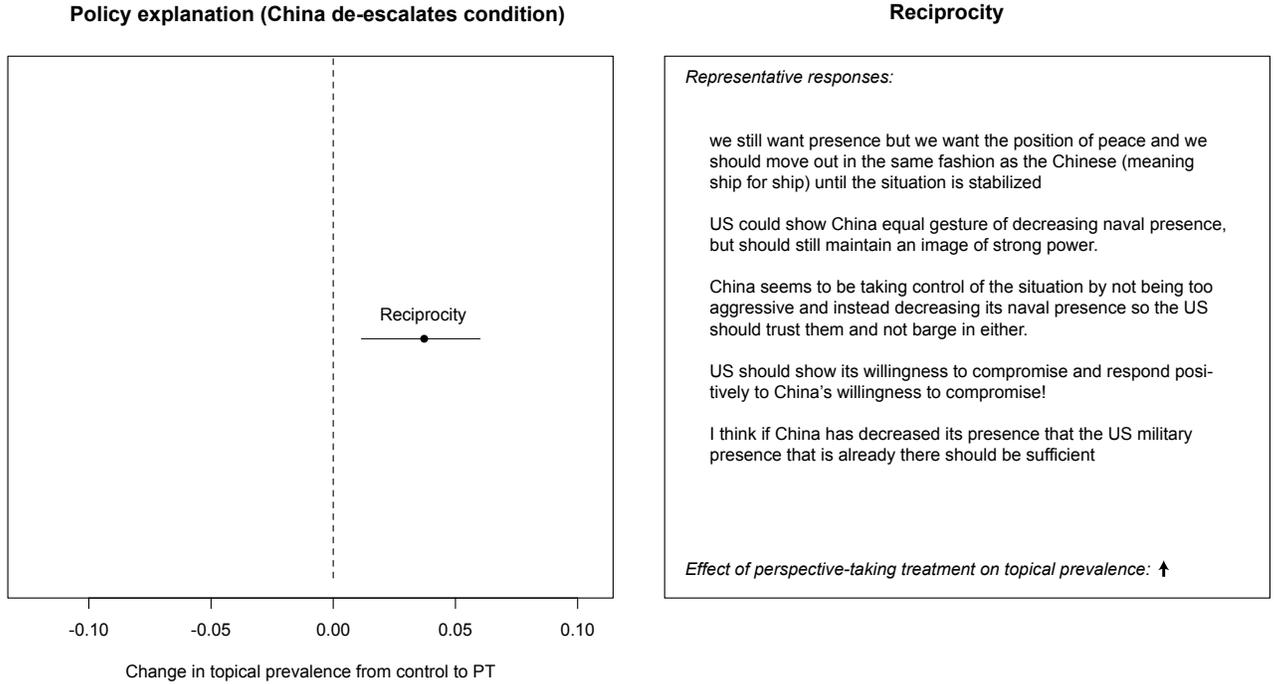
While the technical details of STMs are beyond the scope of this paper (see [Roberts et al. \(2014\)](#) for an introduction), they can be thought of as an automated text analysis technique that models text as a mixture of interpretable "topics." An advantage of STMs are that they are unsupervised, which means they "discover" topics in the text, rather than assuming their existence, which limits the potential for the researcher's prior expectations to influence the topics that are identified. Most importantly for our analysis, unlike more traditional models used in text analysis, such as Latent Dirichlet Allocation (LDA), STMs allow us to leverage information about respondents (in our case, the treatment group to which they were randomly assigned) when structuring the topics, rather than assuming that the topics and topical prevalence are constant across all respondents.

We estimate a series of STMs on our respondents' open-ended explanations for their policy choice, using the models to identify topics within the responses and, most importantly, to measure whether those topics systematically differ in prevalence across the perspective taking treatment and control groups. Since we found the effects of perspective-taking vary depending on whether respondents were in the escalation or de-escalation treatment, we run separate STM models for each escalation treatment group, and then estimate the effect of the perspective-taking treatment within each group.¹⁵

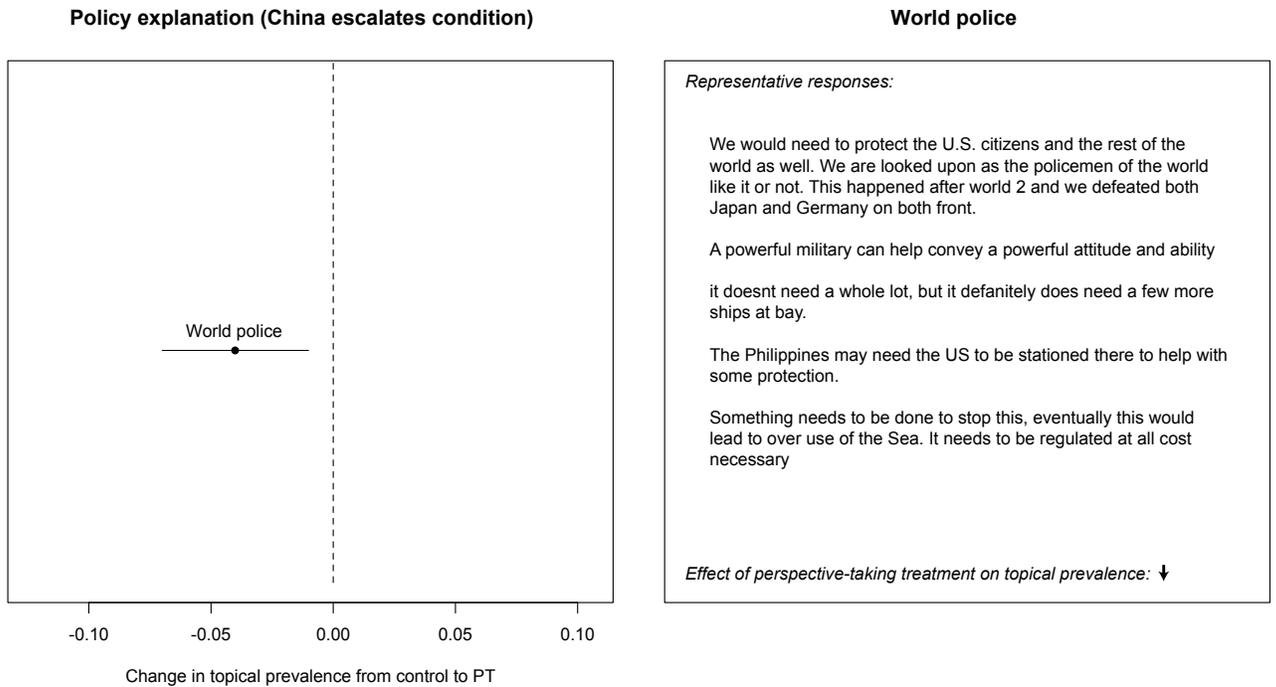
¹⁵For our US respondents, we pre-process the text using the `tm` package in `R`, whereas for our Chinese respondents, we use `Jieba`, a Python-based Chinese word segmentation module. Once the models have been estimated, we then interpret the English-language versions of the topics, as translated by three native Chinese speakers to ensure intercoder reliability.

Figure 3: Structural topic model results: United States

(a) China de-escalates



(b) China escalates



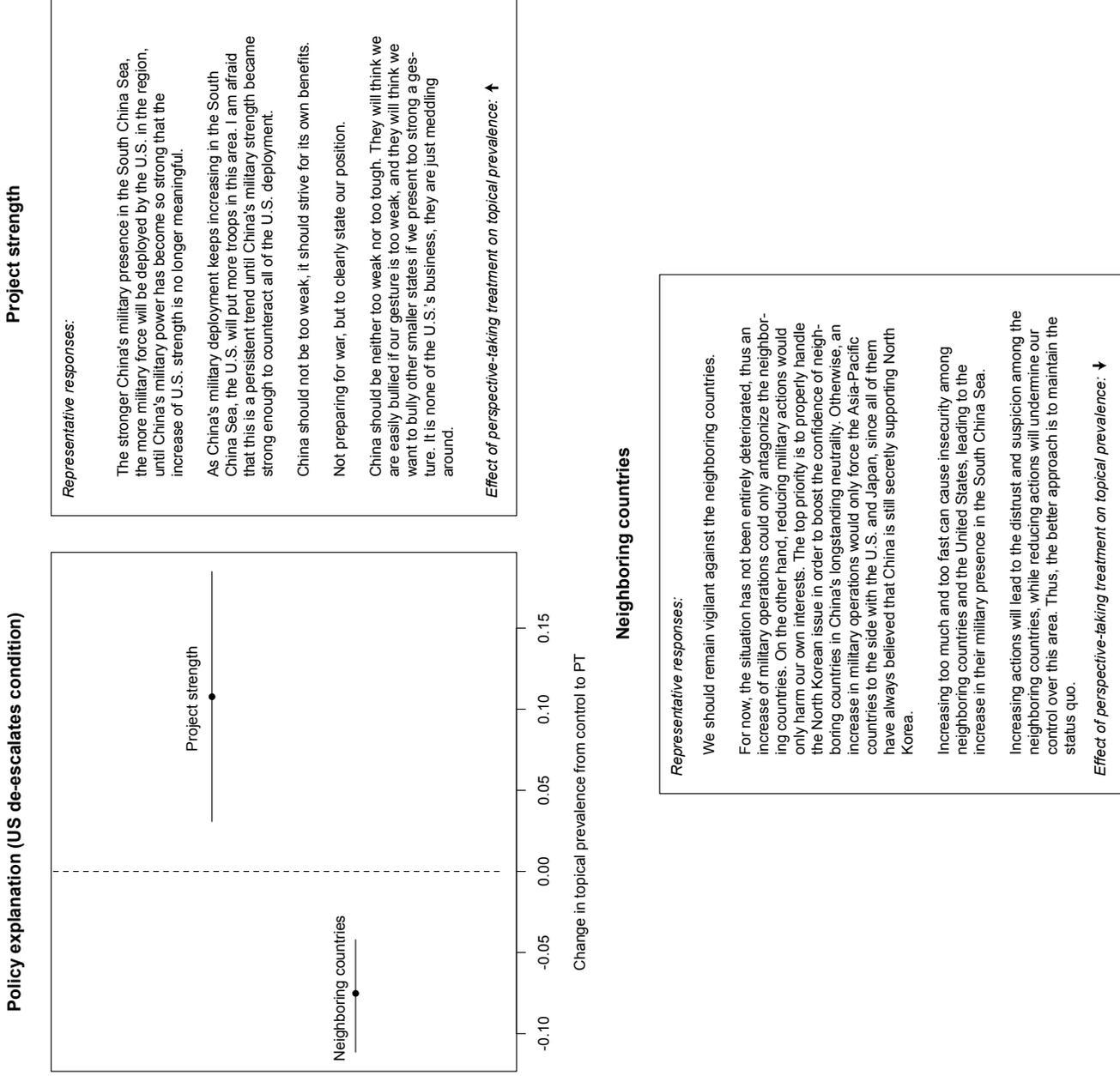
Results from structural topic models, US sample. The left-hand panels display the effect of the perspective-taking treatment on the free responses provided by respondents justifying their policy recommendation; the effect estimates are measures of changes in topical prevalence. The right-hand panels display representative responses for each topic.

For the American sample in Figure 3, we find that the perspective-taking treatment has a significant effect on one topic within each escalation/de-escalation condition. Among participants in the de-escalation condition (Figure 3a), the perspective taking treatment results in a higher usage of language emphasizing reciprocity. As the figure shows, respondents are much more likely to invoke language embracing reciprocal de-escalation, with one respondent writing that the US should “respond positively to China’s willingness to compromise” and another noting that since China decreased its presence the “US could show China’s equal gesture of decreasing naval presence”. Consistent with the above analysis, these findings suggest that the perspective taking treatment makes it more likely that American respondents interpret China’s decrease in military activities as signalling willingness to compromise, which leads them to explain their policy preferences in terms of reciprocal efforts to de-escalate the situation.

Our STM model for the effect of the perspective taking treatment in the Chinese escalation condition also points to the importance of threat perception. As Figure 3b shows, when China increases its military activities, the perspective taking treatment is associated with a significant decrease in language emphasizing Americans’ concern for projecting power and protecting other countries. Encouraging our American respondents to take the perspective of the other side appears to make them less likely to perceive Chinese actions as threatening, which reduces the likelihood that people believe we need to protect ourselves or others from Chinese actions.

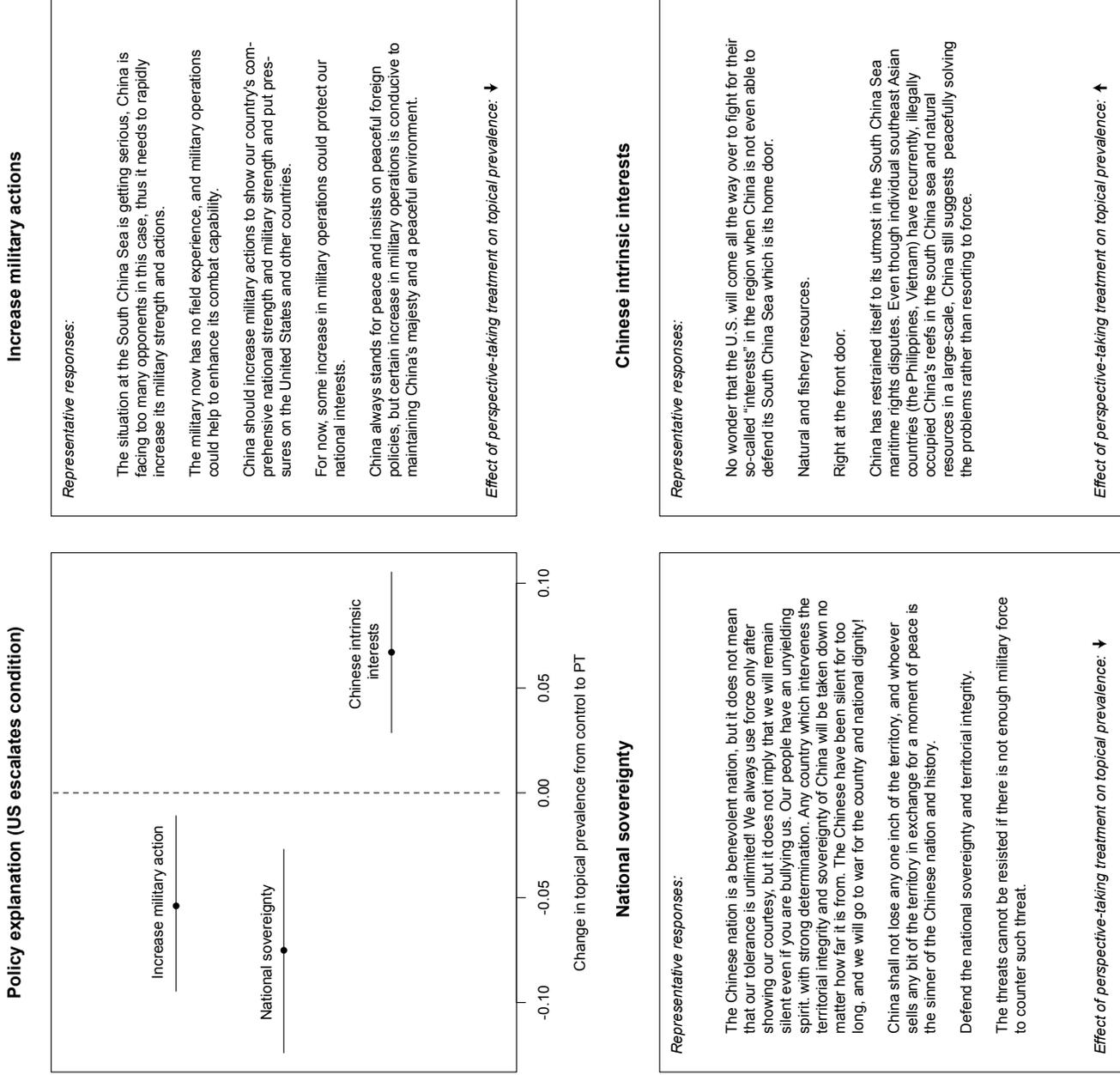
Figures 4-5 display the STM results for our Chinese sample, with the results in the US de-escalation condition displayed in Figure 4, and the results in the US escalation condition displayed in Figure 5. Beginning with the de-escalation condition in Figure 4, we find the perspective-taking treatment significantly affects Chinese participants’ policy explanations in two different ways. First, it increases their emphasis on projecting strength. “The stronger China’s military presence in the South China Sea, " one respondent writes, “the more military force will be deployed by the U.S. in the region, until China’s military power has become so strong that the increase of U.S. strength is no longer meaningful." Thus, another respondent writes, “China should be neither too weak nor too tough. They will think we are easily bullied if our gesture is too weak, and they will think we want to bully other smaller states if we present too strong a gesture. It is none of the U.S.’s business, they are just meddling around." The moderate escalation Chinese respondents display when responding to a US de-escalation in the perspective-taking condition reflects this logic. Second, and especially interestingly, asking participants to assess the conclusions the United States would draw leads them

Figure 4: Structural topic model results: China (United States de-escalates)



Results from structural topic models, China sample (US de-escalates condition). The top-left panel displays the effect of the perspective-taking treatment on the free responses provided by respondents justifying their policy recommendation; the effect estimates are measures of changes in topical prevalence. The remaining panels display representative responses for each topic, translated into English.

Figure 5: Structural topic model results: China (United States escalates)



Results from structural topic models, China sample (US escalates condition). The top-left panel displays the effect of the perspective-taking treatment on the free responses provided by respondents justifying their policy recommendation; the effect estimates are measures of changes in topical prevalence. The remaining panels display representative responses for each topic, translated into English.

to be less likely to think about the strategic calculations of others in the region: participants in the control condition were significantly more likely to discuss the ramifications of China's behavior on its neighboring countries (offering responses like "Increasing actions will lead to the distrust and suspicion among the neighboring countries, while reducing actions will undermine our control over this area. Thus, the better approach is to maintain the status quo."). This is helpful in reminding us that although we think about perspective-taking in a dyadic manner, in a regional context with multiple actors, investing cognitive resources in thinking about one actor's perspective can crowd out the perspective of others.

In the escalation condition in Figure 5, we find that the perspective-taking treatment significantly affects Chinese participants' policy explanations along three different dimensions. First, the perspective-taking treatment decreases the extent to which participants emphasize increased military action. When faced with US escalation, respondents in the perspective-taking treatment are less hawkish than respondents in the control (and thus less likely to offer statements like "The military has no field experience, and military operations could help to enhance its combat capability", or "The situation in the South China Sea is getting serious, China is facing too many opponents in this case, thus it needs to rapidly increase its military strength and actions.") Second, respondents in the perspective-taking treatment are less likely to offer nationalist arguments emphasizing the inviolability of Chinese sovereignty, and thus less likely to make claims like "China shall not lose any one inch of the territory, and whoever sells any bit of the territory in exchange for a moment of peace is the sinner of the Chinese nation and history", or "The Chinese have been silent for too long, and we will go to war for the country and national dignity!" Finally, respondents in the perspective-taking condition are more likely to mention Chinese intrinsic interests, repeatedly mentioning the extent to which the conflict is taking place at China's "front door", or invoking its maritime interests in the region. This suggests that one reason why Chinese respondents de-escalate in the perspective-taking condition when faced with US escalation is because perspective-taking causes them to perceive Chinese interests as sufficiently and self-evidently high that such costly signals are unnecessary.

In sum, perspective-taking affects both *what* respondents want their countries to do, and *why*. Against the palliative effects hypothesis, however, its effects are not always positive, but heavily contingent on the nature of the strategic environment. It makes American respondents place a heavier emphasis on reciprocity and lowers their threat perception, but it also makes Chinese respondents

more interested in projecting strength (and think less about the effect it will have on neighboring countries) when the US shows signs of weakness. When the US escalates, perspective-taking makes Chinese respondents use less hawkish and nationalist language, but also emphasize their interests at stake. In Appendix §5, we explore the origins of these cross-national differences in considerable detail, showing that although there are compositional differences between the two samples (Chinese respondents are both more hawkish and more nationalist, for example, than their American counterparts), these differences are unlikely to explain the patterns of results we see. Similarly, although Chinese and American respondents systematically differ in the images they have of one another (Herrmann and Fischerkeller, 1995), these relational differences also do not appear to be driving the results.

Instead, we attribute the differences displayed above to the situational features of Chinese-US disputes in the South China Sea. While many Americans view the South China Sea as distant and relatively unimportant (as one American respondent put it, “it isn’t our area to dictate to the world who should or shouldn’t have jurisdiction there”), the Chinese perceive it as part of their own backyard (as one Chinese respondent put it, “China and the United States are located on two sides of the planet. On what grounds does the United States have to come all the way over and intervene in our territory?”). In light of these divergent stakes, Chinese and American escalation in the region carries very different meanings to each audience. Since the Americans see little interest in taking the territory, the strategic goal of US presence is simply maintaining the status quo. From this perspective, Chinese de-escalation can be interpreted as a signal of reassurance.

For the Chinese, however, who want to avoid a costly conflict with the United States, but who also perceive the South China Sea as rightfully part of their territory, American de-escalation sends a different signal. As one Chinese respondent wrote, “The fact that the United States decided to withdraw from the South China Sea issue shows that our country has really become powerful. The United States is no longer the hegemonic power that can intervene with any country and any affair in the world, by whatever measures they want to take! I am really proud to be born as a Chinese!” In this sense, the heterogeneous responses of the perspective-taking treatment in China in response to the US escalating and de-escalating suggests a pragmatic response to two very different signals about American resolve. In a series of interviews with Chinese observers, one noted the traditional Chinese military strategy is “敌退我进; 敌进我退”, or “advance when the enemy retreats; retreat when the enemy advances”, the pattern of results we see here. Thus, “the Chinese tend to see U.S.

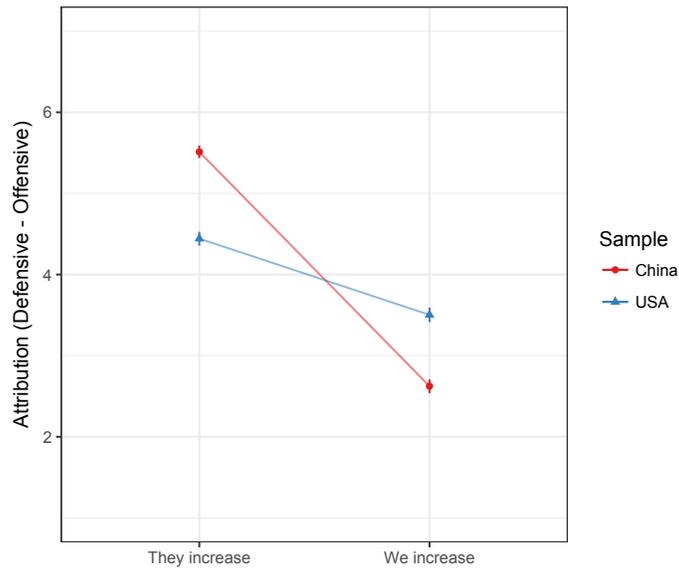
de-escalation as a sign of its weakness; and U.S. escalation, a sign of strength and confidence." Another observer reiterated that the Chinese public perspective on the conflict is influenced by "the official interpretation of the history since 1840" and "typical slogans" that include "落后就要挨打" (we are bullied if we are weak) and "弱肉强食" (the weak are the prey of the strong). "The majority still regards the United States as the strongest power. Thus, the United States can do whatever it wants . . . without constraint." At the same time, "many people do not see the United States as a legitimate player in the game: we are defending the legitimate sovereignty rights over the islands, so we are purely defensive and not a threat to the region. . . the only reason for the United States to stay in the region and the dispute, in particular, is to contain China." In this sense, although both sides see their actions in the region as defensively motivated, they both do so in a very different strategic context.

3.2 Asymmetrical attributions and security dilemma thinking

Thus far, we have examined the effects of perspective-taking on respondents' policy preferences in a security dilemma scenario, finding that IR scholars have perhaps expressed an overly rosy view of empathy's effects on policy preferences. If the situational interpretation above is correct, and the divergent effects of strategic empathy in each country is due to the divergent stakes of the dispute for each side, we should expect to find a similar asymmetry in the attributions each side makes for each other's behavior. Even going beyond the specifics of the US-China case, the security dilemma more broadly is a consequence of asymmetric attributions, as each side perceives their own behavior as defensively-motivated, and the other side's behavior as offensively-motivated. In the context of escalation in the South China Sea, how large is the magnitude of this attribution asymmetry, and can it be mitigated by strategic empathy and perspective-taking?

In a follow-up module, we therefore asked participants in our study to assess the motivations for a hypothetical escalation by each side, on a scale ranging from "For purely defensive reasons" to "For purely offensive reasons." The results, presented in Figure 6, reveals a significant attribution asymmetry among both Americans and Chinese. In each case we find that respondents are much more likely to attribute their country's increased military activities to defensive motivations and the increased military activities of the other country to offensive motivations. Interestingly, however, the magnitude of attribution asymmetry is about three times greater in China than the US, with the attribution asymmetry being 2.89 points ($p < 0.001$) among Chinese and 0.94 points ($p < 0.001$)

Figure 6: Respondents display the attribution asymmetry consistent with security dilemma thinking



Note: Figure 6 displays the average attribution made by Chinese and American respondents for an escalation by the other side (“They increase”) and by their own side (“We increase”). If the attribution made for the other side’s behavior is higher on the scale than the attribution made for one’s own behavior, then the results display an “attribution asymmetry”, an asymmetry in the motivation attributed to an identical action carried out by both sides. The steeper the slope of the line connecting the two points, the stronger the asymmetry. Thus, the figure shows that while both sides tend to see their own actions as more defensive and the other side’s actions as more offensive, the Chinese public displays a much stronger attribution asymmetry than the American public. The figure pools across treatments; in the supplementary analysis we also condition on both the perspective taking treatment and the original escalation treatment, although the results remain the same.

among Americans on the seven point offensive-defensive attribution scale. These findings are not only consistent with security dilemma thinking, but especially on the Chinese side suggest that public opinion would more likely be an accelerator of a potential spiral model than an inhibitor against it. Audiences in both countries believe their nation is acting defensively, but both also believe the other side is acting offensively, likely fostering greater fears among domestic audiences and increased support for tit-for-tat escalation.

Moreover, as we show in Appendix §4.3, when we regress the within-subject measure of attribution asymmetry on participants' predispositional levels of empathy along with a variety of control variables, more empathetic respondents tend to display *larger* degrees of attribution asymmetry, even when controlling for dispositional traits like nationalism or hawkishness. Thus, although some scholars argue that empathy and perspective taking can help alleviate misperceptions and bolster cooperation in international crises, we find that strategic empathy may even exacerbate it.

4 Conclusion

The rise of China has led to an outgrowth of interest in security dilemmas, and a surge of scholarship exploring the ways that tensions and misperceptions can be mitigated. In this paper, we explore the role of strategic empathy or perspective-taking in this process. Given continued tensions in the South China Sea, and concerns about popular nationalism in both the American and Chinese publics, understanding how perceptions on both sides are formed — and shaped by strategic empathy — is a useful endeavor for the study of US-China relations.

Fielding a unique cross-national experiment that examines the security relationship between the US and China in a dyadic manner, we find micro-level evidence consistent with security dilemma thinking in both countries' publics, and that these attribution asymmetries are especially pronounced in China. Despite the volume of literature on China's rise, and the extent to which IR scholars frequently build their arguments on assumptions about the nationalist content of Chinese public opinion, surprisingly little of this literature actually measures this directly.¹⁶ Our findings — consistent with a "liberal pessimist" view of US-China relations (Friedberg, 2005) — thus shed new light on public opinion dynamics in the region, a crucial subject in an era when policymakers are increasingly concerned about the extent to which the mass public in both countries can function as an accelerant

¹⁶For a similar critique, see Johnston (2017).

rather than an inhibitor of great power conflict. Our results show that security dilemma thinking is not merely the preserve of elite decision-makers, but is also detectable in the public writ-large. Indeed, the presence of a substantively large attribution asymmetry in both countries shows that each public is likely both to fear increased military activities by the other side and to believe that they are only trying to preserve the status quo when their country responds in kind. These findings highlight the potential risks between the US and China in the South China sea, and the divergent manners in which citizens in each country understand military signals in the region.

Our study also suggests some skepticism about the palliative effects of empathy hypothesis often adopted by IR scholars. IR scholars have tended to think about strategic empathy's effects in relatively liberal ways, consistent with a broader ethos in IR theory that presumes the existence of a natural harmony of interests, such that conflict is merely the result of misperceptions or bargaining failures (Kertzer, 2016, 41). If only actors are given more information, or decision-makers' excesses can be constrained by a liberal-minded publics, cooperation will be more likely to occur. Our findings push back against both claims. Perspective-taking may reveal the existence of complimentary interests, but it can highlight conflicts of interest as well. We find that experimentally inducing perspective taking makes Americans more likely to recognize conciliatory signals and increase their desire to respond with reciprocal policies of de-escalation, but that Chinese respondents have a very different reaction, causing them to seek to project strength and think less about the regional ramifications when the US de-escalates, but express less nationalist language and back down when the US escalates. We obtain similarly discouraging evidence when we examine natural variation in dispositional levels of empathy among our respondents, as more empathetic respondents display larger attribution asymmetries than their less-empathetic counterparts.

In demonstrating that empathy in IR is not an unvarnished good, these findings suggest several avenues for future research. First, our focus here was on perspective-taking, a cognitive form of empathy, rather than affective forms of empathy that emphasize feeling what others are feeling, as is often emphasized in the existing literature. Future work should test whether there are differential effects between the two. Second, we largely side-step here the question of empathic accuracy: our analysis focuses on what happens when we encourage Chinese and Americans to step into each other's shoes, rather than assess how competently they do so. Future scholarship can benefit from investigating the accuracy of these second-order beliefs directly. Third, it is worth noting that our perspective-taking treatment was deliberately content-free, seeking to induce perspective-

taking rather than priming participants to attribute particular beliefs to the other side. In this sense, the study differs from the types of perspective-taking interventions sometimes used in studies seeking to reduce intergroup prejudice (e.g. Paluck, 2010). Other work can benefit from employing alternate forms of perspective-taking treatments, determining what proportion of the effects come from perspective-taking itself rather than the informational primes. In this manner, by showing that empathy is a broader construct — and thus has more complex effects — than many IR scholars assume, this study points to a broader research agenda on the dynamics of strategic empathy and perspective-taking in international affairs.

References

- Adler, Emanuel and Michael Barnett. 1998. *Security Communities*. Cambridge: Cambridge University Press.
- Aldrich, John H., Christopher Gelpi, Peter Feaver, Jason Reifler and Kristin Thompson Sharp. 2006. “Foreign Policy and the Electoral Connection.” *Annual Review of Political Science* 9:477–502.
- Allison, Graham. 2017. *Destined for War: Can America and China Escape Thucydides’ Trap?* New York, NY: Houghton Mifflin Harcourt.
- Ames, Daniel R. 2004. “Inside the mind reader’s tool kit: projection and stereotyping in mental state inference.” *Journal of personality and social psychology* 87(3):340.
- Atkinson, Carol. 2010. “Does Soft Power Matter? A Comparative Analysis of Student Exchange Programs 1980-2006.” *Foreign Policy Analysis* 6(1):1–22.
- Axelrod, Robert and Robert O. Keohane. 1985. “Achieving Cooperation Under Anarchy: Strategies and Institutions.” *World Politics* 38(1):58–79.
- Baron-Cohen, Simon and Sally Wheelwright. 2004. “The empathy quotient: an investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences.” *Journal of autism and developmental disorders* 34(2):163–175.
- Batson, C. Daniel and Nadia Y. Ahmad. 2009. “Using Empathy to Improve Intergroup Attitudes and Relations.” *Social Issues and Policy Review* 3(1):141–177.
- Bell, Mark S. and Kai Quek. 2018. “Authoritarian Public Opinion and the Democratic Peace.” *International Organization* 72(1):227–242.
- Blight, James G. and Janet M. Lang. 2004. “Lesson Number One: “Empathize With Your Enemy”.” *Peace and Conflict: Journal of Peace Psychology* 10(4):349–368.
- Booth, Ken and Nicholas J. Wheeler. 2008. *The Security Dilemma: Fear, Cooperation and Trust in World Politics*. Basingstoke, UK: Palgrave Macmillan.
- Bronfenbrenner, Urie. 1961. “The Mirror Image in Soviet-American Relations: A Social Psychologist’s Report.” *Journal of Social Issues* 17(3):45–56.
- Brutger, Ryan and Joshua D. Kertzer. 2018. “A Dispositional Theory of Reputation Costs.” *International Organization* 72(3):693–724.
- Butterfield, Herbert. 1951. “The Tragic Element in Modern International Conflict.” *The Review of Politics* 12(2):147–164.
- Caruso, Eugene M., Nicholas Epley and Max H. Bazerman. 2006. “Reactive egoism and unwarranted distrust.” Unpublished manuscript, Harvard University.
- Chang, Linchiat and Jon A. Krosnick. 2010. “Comparing Oral Interviewing with Self-Administered Computerized Questionnaires: An Experiment.” *Public Opinion Quarterly* 74(1):154–167.
- Chen, Dingding, Xiaoyu Pu and Alastair Iain Johnston. 2014. “Debating China’s Assertiveness.” *International Security* 38(3):176–183.

- Chen, Jidong, Jennifer Pan and Yiqing Xu. 2016. "Sources of Authoritarian Responsiveness: A Field Experiment in China." *American Journal of Political Science* 60(2):383–400.
- Christensen, Thomas J. 1999. "China, the U.S.-Japan Alliance, and the Security Dilemma in East Asia." *International Security* 23(4):49–80.
- Christensen, Thomas J. 2002. "The Contemporary Security Dilemma: Deterring a Taiwan Conflict." *The Washington Quarterly* 25(4):7–21.
- Christensen, Thomas J. 2006. "Fostering Stability or Creating a Monster? The Rise of China and U.S. Policy toward East Asia." *International Security* 31(1):81–126.
- Copeland, Dale C. 2000. "The Constructivist Challenge to Structural Realism: A Review Essay." *International Security* 25(2):187–212.
- Crawford, Neta C. 2014. "Institutionalizing passion in world politics: fear and empathy." *International Theory* 6(3):535–557.
- Davis, Mark H. 1980. "A multidimensional approach to individual differences in empathy." *Catalog of Selected Documents in Psychology* 10:85.
- Davis, Mark H. 1983. "Measuring individual differences in empathy: Evidence for a multidimensional approach." *Journal of personality and social psychology* 44(1):113–126.
- Decety, Jean. 2005. Perspective Taking as the Royal Avenue to Empathy. In *Other Minds: How Humans Bridge the Divide Between Self and Others*, ed. Bertram F. Malle and Sara D. Hodges. Guilford Press pp. 143–157.
- Deutsch, Karl W., Sidney A. Burrell and Robert A. Kann. 1957. *Political Community and the North Atlantic Area: International Organization in the Light of Historical Experience*. Princeton, NJ: Princeton University Press.
- Distelhorst, Greg and Yue Hou. 2017. "Constituency Service under Nondemocratic Rule: Evidence from China." *Journal of Politics* 79(3):1024–1040.
- Epley, Nicholas, Boaz Keysar, Leaf Van Boven and Thomas Gilovich. 2004. "Perspective Taking as Egocentric Anchoring and Adjustment." *Journal of Personality and Social Psychology* 87(3):327–339.
- Epley, Nicholas and Eugene M. Caruso. 2008. Perspective Taking: Misstepping Into Others' Shoes. In *The handbook of imagination and mental simulation*, ed. Keith D. Markman, William M.P. Klein and Julie A. Suhr. New York: Psychology Press pp. 295–309.
- Epley, Nicholas, Eugene M. Caruso and Max H. Bazerman. 2006. "When Perspective Taking Increases Taking: Reactive Egoism in Social Interaction." *Journal of Personality and Social Psychology* 91(5):872–889.
- Epley, Nicholas and Thomas Gilovich. 2006. "The anchoring-and-adjustment heuristic: Why the adjustments are insufficient." *Psychological Science* 17(4):311–318.
- Fearon, James. 1995. "Rationalist Explanations for War." *International Organization* 49(3):379–414.
- Fearon, James D. 1994. "Domestic Political Audiences and the Escalation of International Disputes." *American Political Science Review* 88(3):577–592.
- Feldman, Stanley and Marco R. Steenbergen. 2001. "The Humanitarian Foundation of Public Support for Social Welfare." *American Journal of Political Science* 45(3):658–677.
- Finnemore, Martha and Kathryn Sikkink. 1998. "International Norm Dynamics and Political Change." *International Organization* 52(4):887–917.
- Freedman, Lawrence. 2013. *Strategy: a history*. Oxford: Oxford University Press.
- Friedberg, Aaron L. 2005. "The Future of U.S.-China Relations: Is Conflict Inevitable?" *International Security* 30(2):7–45.
- Gelpi, Christopher and Joseph M. Grieco. 2015. "Competency Costs in Foreign Affairs: Presidential Performance in International Conflicts and Domestic Legislative Success, 1953–2001." *American Journal of Political Science* 59(2):440–456.
- Gilpin, Robert. 1981. *War and Change in World Politics*. Cambridge: Cambridge University Press.
- Glaser, Charles L. 1997. "The Security Dilemma Revisited." *World Politics* 50(1):171–201.
- Goldstein, Avery. 2005. *Rising to the Challenge: China's Grand Strategy and International Security*.

- Stanford, CA: Stanford University Press.
- Gries, Peter Hays. 2001. "Tears of Rage: Chinese Nationalist Reactions to the Belgrade Embassy Bombing." *China Journal* 46(1):25–43.
- Gries, Peter Hays. 2005. "Social Psychology and the Identity-Conflict Debate: Is the "China Threat" Inevitable?" *European Journal of International Relations* 11(2):235–265.
- Gross, Kimberly. 2008. "Framing Persuasive Appeals: Episodic and Thematic Framing, Emotional Response, and Policy Opinion." *Political Psychology* 29(2):169–192.
- Head, Naomi. 2016. "Costly encounters of the empathic kind: a typology." *International Theory* 8(1):171–199.
- Herrmann, Richard. 1988. "The Empirical Challenge of the Cognitive Revolution: A Strategy for the Drawing Inferences about Perceptions." *International Studies Quarterly* 32(2):175–203.
- Herrmann, Richard K. and Michael P. Fischerkeller. 1995. "Beyond the Enemy Image and Spiral Model: Cognitive-Strategic Research After the Cold War." *International Organization* 49(3):415–50.
- Herz, John H. 1950. "Idealist internationalism and the security dilemma." *World politics* 2(2):157–180.
- Holmes, Marcus. 2018. *Face-to-Face Diplomacy: Social Neuroscience and International Relations*. Cambridge, UK: Cambridge University Press.
- Holmes, Marcus and Keren Yarhi-Milo. 2017. "The Psychological Logic of Peace Summits: How Empathy Shapes Outcomes of Diplomatic Negotiations." *International Studies Quarterly* 61(1):107–122.
- Holsti, Ole R. 2004. *Public Opinion and American Foreign Policy*. Revised edition ed. Ann Arbor, MI: University of Michigan Press.
- Huang, Haifeng. 2015. "International Knowledge and Domestic Evaluations in a Changing Society: The Case of China." *American Political Science Review* 109(3):613–634.
- Ikenberry, G. John. 2008. "The Rise of China and the Future of the West: Can the Liberal System Survive?" *Foreign Affairs* 87(1):23–37.
- Jervis, Robert. 1976. *Perception and Misperception in International Politics*. Princeton, NJ: Princeton University Press.
- Jervis, Robert. 1978. "Cooperation Under the Security Dilemma." *World Politics* 30(2):167–214.
- Jervis, Robert, Richard Ned Lebow and Janice Gross Stein. 1985. *Psychology and Deterrence*. Baltimore: John Hopkins University Press.
- Johnston, Alastair Iain. 2003. "Is China a Status Quo Power?" *International Security* 27(4):5–56.
- Johnston, Alastair Iain. 2012. "What (If Anything) Does East Asia Tell Us About International Relations Theory?" *Annual Review of Political Science* 15(1):53–78.
- Johnston, Alastair Iain. 2017. "Is Chinese Nationalism Rising: Evidence from Beijing." *International Security* 41(3):7–43.
- Keller, Jonathan W. and Yi Edward Yang. 2009. "Empathy and Strategic Interaction in Crises: A Poliheuristic Perspective." *Foreign Policy Analysis* 5(2):169–189.
- Keohane, Robert O. 1984. *After Hegemony: Cooperation and Discord in the World Political Economy*. Princeton, NJ: Princeton University Press.
- Kertzer, Joshua D. 2016. *Resolve in International Politics*. Princeton, NJ: Princeton University Press.
- Kertzer, Joshua D. and Brian C. Rathbun. 2015. "Fair is Fair: Social Preferences and Reciprocity in International Politics." *World Politics* 67(4):613–655.
- Kertzer, Joshua D. and Jonathan Renshon. 2014. "Putting Things in Perspective: Perspective-Taking in Experimental Political Science." Working paper.
- 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):234–249.
- Kinder, Donald R and Cindy D Kam. 2010. *Us against them: Ethnocentric foundations of American opinion*. Chicago, IL: University of Chicago Press.

- Kugler, Jacek and Douglas Lemke. 2000. The Power Transition Research Program: Assessing Theoretical and Empirical Advances. In *Handbook of War Studies II*, ed. Manus I. Midlarsky. Ann Arbor: University of Michigan Press pp. 129–163.
- Larson, Deborah Welch. 1994. “The Role of Belief Systems and Schemas in Foreign Policy Decision-Making.” *Political Psychology* 15(1):17–33.
- Levy, Jack S. 1987. “Declining power and the preventive motivation for war.” *World Politics* 40(1):82–107.
- Li, Xiaojun, Jianwei Wang and Dingding Chen. 2016. “Chinese Citizens’ Trust in Japan and South Korea: Findings from a Four-City Survey.” *International Studies Quarterly* 60(4):778–789.
- Liff, Adam P. and G. John Ikenberry. 2014. “Racing toward Tragedy?: China’s Rise, Military Competition in the Asia Pacific, and the Security Dilemma.” *International Security* 39(2):52–91.
- Ludwig, Andrew. 2017. “In the Absence of American Policy, a Security Dilemma Grows in Asia.” China US Focus.
- Marinov, Nikolay, William G. Nomikos and Josh Robbins. 2015. “Does Electoral Proximity Affect Security Policy?” *Journal of Politics* 77(3):762–773.
- Mearsheimer, John J. 2001. *The Tragedy of Great Power Politics*. New York: W.W. Norton.
- Mearsheimer, John J. 2014. “Can China Rise Peacefully?” *National Interest* .
URL: <http://nationalinterest.org/commentary/can-china-rise-peacefully-10204>
- Medeiros, Evan S and M Taylor Fravel. 2003. “China’s new diplomacy.” *Foreign Affairs* 82(6):22–35.
- Meng, Tianguang, Jennifer Pan and Yang Ping. 2017. “Conditional Receptivity to Citizen Participation: Evidence from a Survey Experiment in China.” *Comparative Political Studies* 50(4):399–433.
- Mitzen, Jennifer and Randall L. Schweller. 2011. “Knowing the Unknown Unknowns: Misplaced Certainty and the Onset of War.” *Security Studies* 20(1):2–35.
- Morgenthau, Hans J. 1985. *Politics Among Nations: The Struggle for Power and Peace*. Brief ed. Boston, MA: McGraw-Hill.
- Morrow, James D. 1994. *Game Theory for Political Scientists*. Princeton, NJ: Princeton University Press.
- Newman, Benjamin J., Todd K. Hartman, Patrick L. Lown and Stanley Feldman. 2015. “Easing the Heavy Hand: Humanitarian Concern, Empathy, and Opinion on Immigration.” *British Journal of Political Science* 45(3):583–607.
- Pagano, Sabrina J. and Yuen J. Huo. 2007. “The Role of Moral Emotions in Predicting Support for Political Actions in Post-War Iraq.” *Political Psychology* 28(2):227–255.
- Paluck, Elizabeth Levy. 2010. “Is it Better Not to Talk? Group Polarization, Extended Contact, and Perspective Taking in Eastern Democratic Republic of Congo.” *Personality and Social Psychology Bulletin* 36(9):1170–1185.
- Pettigrew, Thomas F. 1998. “Intergroup Contact Theory.” *Annual Review of Psychology* 49:65–85.
- Pierce, Jason R., Gavin J. Kilduff, Adam D. Galinsky and Niro Sivanathan. 2013. “From Glue to Gasoline: How Competition Turns Perspective Takers Unethical.” *Psychological Science* 24(10):1986–1994.
- Quek, Kai. 2017. “Type II Audience Costs.” *Journal of Politics* 79(4):1438–1443.
- Reilly, James. 2011. *Strong Society, Smart State: The Rise of Public Opinion in China’s Japan Policy*. New York, NY: Columbia University Press.
- Roberts, Margaret E., Brandon M. Stewart, Dustin Tingley, Christopher Lucas, Jetson Leder-Luis, Shana Kushner Gadarian, Bethany Albertson and David G. Rand. 2014. “Structural Topic Models for Open-Ended Survey Responses.” *American Journal of Political Science* 58(4):1064–1082.
- Ross, Lee. 1977. The intuitive psychologist and his shortcomings: Distortions in the attribution process. In *Advances in Experimental Social Psychology, Vol. 10*, ed. Leonard Berkowitz. New York: Academic Press pp. 173–220.
- Ross, Robert S. 2009. “China’s Naval Nationalism: Sources, Prospects, and the U.S. Response.” *International Security* 34(2):46–81.
- Sagan, Scott D. and Benjamin A. Valentino. 2017. “Revisiting Hiroshima in Iran: What Americans

- Really Think about Using Nuclear Weapons and Killing Noncombatants.” *International Security* 42(1):41–79.
- Sassenrath, Claudia, Sara D. Hodges and Stefan Pfattheicher. 2016. “It’s All About the Self: When Perspective Taking Backfires.” *Current Directions in Psychological Science* 25(6):405–410.
- Saunders, Elizabeth N. 2015. “War and the Inner Circle: Democratic Elites and the Politics of Using Force.” *Security Studies* 24(3):466–501.
- Schelling, Thomas C. 1960. *The Strategy of Conflict*. Cambridge, MA: Harvard University Press.
- Schweller, Randall L. and Xiaoyu Pu. 2011. “After Unipolarity: China’s Visions of International Order in an Era of U.S. Decline.” *International Security* 36(1):41–72.
- Shore, Zachary. 2014. *A Sense of the Enemy: The High Stakes History of Reading Your Rival’s Mind*. New York, NY: Oxford University Press.
- Simonovits, Gábor, Gábor Kézdi and Péter Kardos. 2018. “Seeing the World Through the Other’s Eye: An Online Intervention Reducing Ethnic Prejudice.” *American Political Science Review* 112(1):186–193.
- Singer, Tania and Ernst Fehr. 2005. “The Neuroeconomics of Mind Reading and Empathy.” *American Economic Review* 95(2):340–345.
- Smith, M. Brewster. 2004. “Realistic Empathy: A Key to Sensible International Relations.” *Peace and Conflict: Journal of Peace Psychology* 10(4):335–339.
- Stein, Janice Gross. 1988. “Building Politics into Psychology: The Misperception of Threat.” *Political Psychology* 9(2):245–271.
- Steinberg, James and Michael E. O’Hanlon. 2014. *Strategic Reassurance and Resolve: U.S.-China Relations in the Twenty-First Century*. Princeton, NJ: Princeton University Press.
- Tang, Shiping. 2009. “The Security Dilemma: A Conceptual Analysis.” *Security Studies* 18(3):587–623.
- Tarrant, Mark, Raff Calitri and Dale Weston. 2012. “Social identification Structures the Effects of Perspective Taking.” *Psychological Science* 23(9):973–978.
- Teng, Jianqun. 2017. “The Sino-US Security Dilemma: The Root Cause and Way Out.” China Institute of International Studies.
- Tingley, Dustin. 2017. “Rising Power on the Mind.” *International Organization* 71(S1):S165–S188.
- Tomz, Michael. 2007. “Domestic Audience Costs in International Relations: An Experimental Approach.” *International Organization* 61(4):821–40.
- Tomz, Michael R. and Jessica L. P. Weeks. 2013. “Public Opinion and the Democratic Peace.” *American Political Science Review* 107(4):849–865.
- Trager, Robert F. and Lynn Vavreck. 2011. “The Political Costs of Crisis Bargaining: Presidential Rhetoric and the Role of Party.” *American Journal of Political Science* 55(3):526–545.
- Truex, Rory. 2016. *Making Autocracy Work: Representation and Responsiveness in Modern China*. New York, NY: Cambridge University Press.
- Waldman, Matt. 2014. “Strategic Empathy.” New America Foundation Policy Paper.
- Wallace, Jeremy L. and Jessica Chen Weiss. 2015. “The Political Geography of Nationalist Protest in China: Cities and the 2012 Anti-Japanese Protests.” *The China Quarterly* 222:403–429.
- Waltz, Kenneth N. 1979. *Theory of International Politics*. Boston: McGraw-Hill.
- Weiss, Jessica Chen. 2013. “Authoritarian Signaling, Mass Audiences, and Nationalist Protest in China.” *International Organization* 67(1):1–35.
- Weiss, Jessica Chen. 2014. *Powerful Patriots: Nationalist Protest in China’s Foreign Relations*. Oxford: Oxford University Press.
- Weiss, Jessica Chen and Allan Dafoe. 2016. “Authoritarian Audiences and Government Rhetoric in International Crises: Evidence from China.” Working paper.
- Wendt, Alexander. 1999. *Social Theory of International Politics*. Cambridge, UK: Cambridge University Press.
- White, Ralph K. 1986. “Empathy as an intelligence tool.” *International Journal of Intelligence and CounterIntelligence* 1(1):57–75.

STRATEGIC EMPATHY AND THE SECURITY DILEMMA:
Supplementary Appendix

Contents

1	Note on survey partner	2
	Table 1: Sample characteristics: Chinese sample	2
	Table 2: Sample characteristics: US sample	3
2	Translation protocol for survey instrument	4
3	Survey instrument	4
4	Supplementary analysis	8
4.1	Causal average complier effects	8
	Table 3: Causal average complier effect (CACE) estimates, US sample	8
	Table 4: Causal average complier effect (CACE) estimates, Chinese sample	8
4.2	Balance tests	10
	Figure 1: Balance tests: China experiment	11
	Figure 2: Balance tests: US experiment	12
4.3	Exploiting natural variation in empathy	13
	Table 5: The dispositional empathy results offer little support for the palliative effects of empathy hypothesis (I)	14
	Table 6: The dispositional empathy results offer little support for the palliative effects of empathy hypothesis (II)	16
	Table 7: More empathetic respondents display larger attribution asymmetries (I)	17
	Table 8: More empathetic respondents display larger attribution asymmetries (II)	18
5	Exploring cross-national differences	19
5.1	Compositional differences	19
	Figure 3: Compositional differences across samples	23
	Figure 4: Heterogeneous treatment effects across dispositional variables	24
5.2	Relational differences	25
	Figure 5: Relational differences across samples	27
	Figure 6: Heterogeneous treatment effects across image variables	29
	Figure 7: Heterogeneous treatment effects across attribution variables	30
5.3	Situational differences	31

1 Note on survey partner

Political surveys are sensitive in China. This is especially the case for a national-level survey fielded across all provinces and capital municipalities of China; as Table 1 shows, the resulting sample skews closely to demographic targets (as does the US sample - see Table 2). Given the political conditions, our partner company would only agree to field our experiment on the condition of strict confidentiality. We persuaded our partner to make an exception for the review process; however, no identifying information can be revealed to others beyond this exception as the leakage may impact the company's interests in China. (This would not prevent us from making available our anonymized data or complying with the journal's replication policy.)

[Remainder of section is redacted, except for confidential drafts under review.]

While political/social desirability bias can exist in any survey sample, the concern may be greater in China due to the political environment. To manage this concern, we opted for an Internet survey instead of a face-to-face or telephone survey. Researchers have shown that anonymous online surveys reduce desirability bias (Chang and Krosnick, 2010). Indeed, we found that respondents seem to be quite willing to disapprove of the government when it does something that goes against their individual preferences. For example, in a 2015 study, one of the co-authors found that the majority of respondents did not approve of a Chinese leader who had threatened to use force in a territorial dispute but subsequently did not. Of course, even if desirability bias exists in China, it does not make the results less useful. The results remain relevant because the same desirability bias also operates in the real world.

Table 1: Sample characteristics: Chinese sample

Characteristic		2010 National Census		
		Full Population	Age ≥ 18	2016 sample
Gender	Male	51.2%	50.52%	52.6%
Geography	Northern	12.7%	-	13.1%
	Northeastern	8.8%	-	7.3%
	Eastern	30.2%	-	28.0%
	Central & Southern	27.0%	-	25.9%
	Southwestern	14.2%	-	15.9%
	Northwestern	7.2%	-	7.3%
Age	≤ 19	24.1%	4.0%	2.0%
	20-29	17.1%	21.7%	25.6%
	30-39	16.1%	20.4%	24.1%
	40-49	17.3%	21.9%	18.2%
	≥ 50	25.3%	32.0%	30.0%
Ethnicity	Han	91.6%	92.3%	94.9%

Table 2: Sample characteristics: US sample

Characteristic		Sample	Population Target
Gender	Male	0.45	0.49
	Female	0.55	0.51
Age	18-24	0.13	0.13
	25-44	0.33	0.34
	45-64	0.34	0.34
	≥ 65	0.21	0.19
Education	High school or less	0.34	0.42
	Some college/university	0.12	0.19
	College/university	0.37	0.28
	Graduate/professional school	0.17	0.10

2 Translation protocol for survey instrument

The survey instrument was first written in the English language. For our Chinese respondents, the instrument was translated into Chinese and pre-tested in multiple stages to ensure accuracy. First, a native Chinese speaker (C1) translated the English instrument into Chinese. Next, one of the co-authors (A1) vetted the translated text and produced the first pre-test version. Both C1 and A1 are effectively bilingual. The focus of the translation was accuracy of meaning instead of a literal word-to-word translation, which can cause meaning distortions when the translated words combine into sentence form. Subsequently, this survey version is tested separately with four other Chinese speakers (C2, C3, C4, C5). Their feedback was used by A1 to produce the second pre-test version. This version is pre-tested with a small sample of Chinese natives (NS) selected from each of the different regions of China to check if the questions were clear and natural to them. Unclear or awkward wordings (few were detected at this stage) were revised in the final translated version. Sample NS was monolingual and not aware of the study design. Finally, the Chinese-language survey was also pre-tested by four additional Chinese language speakers, one native and three bilingual, to flag any problematic or confusing language.

3 Survey instrument

Introduction

下列问题涉及中国与世界上其他国家的关系。

The following questions are about China's relations with other countries around the world.

最近南海局势备受关注。区域内的许多国家宣称，他们在有争议的国际水域拥有权利，而这些水域有许多自然资源、渔业资源和贸易航道。这些在近年来愈演愈烈的外交争端中极为关键。中国关注美国在该地区的强硬举措。

Recently there has been much attention over tensions in the South China Sea. Multiple countries in the region have claimed rights to disputed international waters, which are home to a wealth of natural resources, fisheries, and trade routes — all of which are at stake in the increasingly frequent diplomatic standoffs. China is concerned about the United States' assertiveness in the region.

Writing Task I

假设，美国决定[增加/减少]在南海战略海域的海军部署。您认为，中国会从美国的行动中得出什么样的结论？

Suppose the U.S. has decided to [increase / decrease] naval deployments in strategic maritime zones in the South China Sea. In your view, what conclusion would China draw from the U.S.'s action? [Writing task I: Open-ended response]

Perspective-taking treatment

背景介绍:

- 南海近来局势紧张。
- 许多区域内国家主张对有争议的水域拥有权利。中国关注美国在区域内的强硬举措
- 假设，美国决定[增加/减少]在南海战略海域的海军部署。

中国国内在讨论，是否应当改变中国在南海的军事存在。

您认为，若中国决定增加其在南海的军事行动，美国会从中得出什么样的结论？

Background:

- There has recently been tension in the South China Sea.
- Multiple countries in the region have claimed rights to disputed waters. China is concerned about the U.S.'s assertiveness in the region.
- Suppose the U.S. has decided to [increase / decrease] naval deployments in strategic maritime zones in the South China Sea.

There is a debate in China about whether China should change their military presence in the South China Sea.

If China were to increase its military activities in the South China Sea, in your view, what conclusion would the U.S. draw from China's action? [Writing task II: Open-ended response]

Policy choice

中国国内在讨论，是否应当改变中国在南海的军事存在。

在这种情况下，您认为中国在南海的军事行动应当增加、减少，还是保持不变？

[增加/减少/保持不变]

There is a debate in China about whether China should change their military presence in the South China Sea.

In this scenario, do you think that China's military activities around the South China Sea should increase, decrease, or be kept about the same? [Increase / decrease / kept about the same]

[If increase]

中国的军事行动应当增加很多，还是增加一些？

[增加很多/增加一些]

Should China's military activities increase by a lot, or only somewhat? [Increase / decrease / kept about the same]

[If decrease]

中国的军事行动应当减少很多，还是减少一些？

[减少很多/减少一些]

Should China's military activities decrease by a lot, or only somewhat? [Increase / decrease / kept about the same]

[If kept about the same]

您是倾向增加军事行动、倾向减少军事行动、还是不倾向任何一方？

[倾向增加/倾向减少/不倾向任何一方]

Do you lean toward increasing military activities, lean toward decreasing them, or don't lean either way? [Increase / decrease / kept about the same]

请简要解释您的回答:

Please briefly explain your response.

Attribution and cooperation module

对于下列说法, 请选择最接近您观点的回答:

- 如果中国增加在南海的军事行动, 您认为那是: [1 (纯粹为了防御) – 7 (纯粹为了进攻)]
- 如果美国增加在南海的军事行动, 您认为那是: [1 (纯粹为了防御) – 7 (纯粹为了进攻)]

For each of the following statements, please select the answer that most closely matches your own views.

- If China were to increase its military activities in the South China Sea, do you think it would be: [1 (Purely for defensive reasons) – 7 (Purely for offensive reasons)]
- If the U.S. were to increase its military activities in the South China Sea, do you think it would be: [1 (Purely for defensive reasons) – 7 (Purely for offensive reasons)]

对于下列说法, 请选择最接近您观点的回答:

- 如果中国增加在南海的军事行动, 您认为这会如何影响中美经济合作的可能? [1 (比较可能合作) – 7 (比较不可能合作)]
- 如果美国增加在南海的军事行动, 您认为这会如何影响中美经济合作的可能? [1 (比较可能合作) – 7 (比较不可能合作)]

For each of the following statements, please select the answer that most closely matches your own views.

- If China were to increase its military activities in the South China Sea, how do you think it would affect the possibility of Sino-American economic cooperation? [1 (More likely to cooperate) – 7 (Less likely to cooperate)]
- If the U.S. were to increase its military activities in the South China Sea, how do you think it would affect the possibility of Sino-American economic cooperation? [1 (More likely to cooperate) – 7 (Less likely to cooperate)]

4 Supplementary analysis

4.1 Causal average complier effects

Table 3: Causal average complier effect (CACE) estimates, US sample

τ	China de-escalates				China escalates			
	Ratio estimator		2SLS estimate		Ratio estimator		2SLS estimate	
	CACE	95% CIs	CACE	95% CIs	CACE	95% CIs	CACE	95% CIs
0	-0.27	(-0.51, -0.05)	-0.27	(-0.5, -0.05)	-0.06	(-0.29, 0.18)	-0.06	(-0.29, 0.18)
5	-0.35	(-0.64, -0.06)	-0.28	(-0.52, -0.05)	-0.07	(-0.36, 0.22)	-0.06	(-0.3, 0.19)
10	-0.36	(-0.67, -0.06)	-0.25	(-0.5, -0.01)	-0.07	(-0.38, 0.23)	-0.07	(-0.33, 0.19)
15	-0.39	(-0.72, -0.07)	-0.26	(-0.53, 0)	-0.08	(-0.4, 0.24)	-0.06	(-0.33, 0.21)
20	-0.42	(-0.77, -0.07)	-0.28	(-0.56, -0.01)	-0.08	(-0.42, 0.26)	-0.02	(-0.31, 0.26)
25	-0.43	(-0.8, -0.08)	-0.33	(-0.62, -0.05)	-0.09	(-0.45, 0.27)	0.02	(-0.28, 0.32)
30	-0.47	(-0.86, -0.08)	-0.38	(-0.69, -0.08)	-0.1	(-0.48, 0.29)	0.01	(-0.31, 0.32)
35	-0.51	(-0.95, -0.09)	-0.43	(-0.75, -0.1)	-0.1	(-0.52, 0.31)	-0.06	(-0.4, 0.27)
40	-0.57	(-1.06, -0.1)	-0.41	(-0.76, -0.06)	-0.11	(-0.56, 0.34)	-0.09	(-0.45, 0.26)
45	-0.61	(-1.13, -0.11)	-0.45	(-0.82, -0.08)	-0.12	(-0.62, 0.37)	-0.03	(-0.4, 0.35)
50	-0.66	(-1.25, -0.12)	-0.43	(-0.82, -0.04)	-0.13	(-0.67, 0.4)	-0.13	(-0.53, 0.27)

Table 3 presents CACE estimates, defining compliance at different percentile thresholds (τ) of response length. The first two columns in each panel uses the ratio estimator $CACE = \frac{ITT}{Pr(Complied)}$ with 95% bootstrapped CIs, while the next two results from a 2SLS model with 95% normal theory CIs. Note that at 0% non-compliance, the CACE estimates replicate the ATEs reported in the main text.

Table 4: Causal average complier effect (CACE) estimates, Chinese sample

	US de-escalates				US escalates			
	Ratio estimator		2SLS estimate		Ratio estimator		2SLS estimate	
	CACE	95% CIs	CACE	95% CIs	CACE	95% CIs	CACE	95% CIs
0	0.23	(0.002, 0.44)	0.23	(0.01, 0.45)	-0.23	(-0.41, -0.04)	-0.23	(-0.42, -0.04)
5	0.24	(0.002, 0.45)	0.25	(0.03, 0.47)	-0.24	(-0.42, -0.04)	-0.18	(-0.38, 0.01)
10	0.24	(0.002, 0.45)	0.25	(0.03, 0.47)	-0.24	(-0.42, -0.04)	-0.18	(-0.38, 0.01)
15	0.25	(0.002, 0.48)	0.38	(0.14, 0.62)	-0.27	(-0.48, -0.05)	-0.04	(-0.26, 0.18)
20	0.25	(0.002, 0.48)	0.38	(0.14, 0.62)	-0.27	(-0.48, -0.05)	-0.04	(-0.26, 0.18)
25	0.28	(0.003, 0.54)	0.46	(0.19, 0.72)	-0.3	(-0.55, -0.06)	0.12	(-0.13, 0.36)
30	0.31	(0.003, 0.59)	0.52	(0.24, 0.81)	-0.34	(-0.61, -0.06)	0.14	(-0.12, 0.4)
35	0.32	(0.003, 0.61)	0.56	(0.27, 0.85)	-0.35	(-0.63, -0.07)	0.08	(-0.2, 0.35)
40	0.36	(0.003, 0.68)	0.47	(0.16, 0.79)	-0.38	(-0.69, -0.07)	0.05	(-0.24, 0.34)
45	0.39	(0.004, 0.74)	0.41	(0.07, 0.74)	-0.42	(-0.76, -0.08)	0.04	(-0.27, 0.35)
50	0.43	(0.004, 0.82)	0.35	(0, 0.71)	-0.49	(-0.89, -0.09)	0.05	(-0.3, 0.39)

Table 3 presents CACE estimates, defining compliance at different percentile thresholds (τ) of response length. The first two columns in each panel uses the ratio estimator $CACE = \frac{ITT}{Pr(Complied)}$ with 95% bootstrapped CIs, while the next two results from a 2SLS model with 95% normal theory CIs. Note that at 0% non-compliance, the CACE estimates replicate the ATEs reported in the main text.

A potential concern when evaluating the effects of perspective-taking is that our experiment is akin to an encouragement design: participants are asked to think through and write out what conclusions the other country would draw from a particular foreign policy behavior. As with any encouragement design, then, it is important to address potential concerns about non-compliance: most

notably, some participants in the treatment condition, who are encouraged to engage in perspective-taking may not take the exercise seriously and refuse to engage in it; in a potential outcomes framework, these would be the “never takers.” An examination of the contents of the second writing task suggest clear evidence that some participants did not comply with the treatment, writing responses in the US survey like “Balls”, “Merica” or “no comment”.

To explore the implications of non-compliance for our interpretation of the experimental results, we combine an instrumental variables approach with sensitivity analyses. First, we calculated the causal average complier effects (CACE) using the [Angrist, Imbens and Rubin \(1996\)](#) ratio estimator, in which the CACE is the intention-to-treat (ITT) effect divided by the proportion of compliers in the treatment condition. To avoid making subjective coding decisions about the types of responses that count as complying, we define compliance using a sensitivity analysis approach, defining compliance thresholds at different percentiles of response length to the second writing task completed by participants in the perspective-taking treatment, under the assumption that respondents who wrote longer responses are more likely to have complied with the treatment. We calculate response length in the US survey by simply calculating the number of characters of the free response; for the Chinese survey we first pre-process the free-responses using [Jieba](#), a Python-based Chinese word segmentation module. The ratio estimator results are presented in the first two columns in each panel in [Table 3](#) for the US survey, and in [Table 4](#) for the Chinese survey, with 95% bootstrapped confidence intervals. We also estimate the CACE using a more formal 2-stage least squares (2SLS) approach, instrumenting for compliance with the treatment using random assignment to the perspective-taking treatment ([Gerber and Green, 2012](#)); these results are shown in the third and fourth columns in each panel of each table, with 95% normal theory confidence intervals.

Despite the crudeness of our indicators of compliance, we find results largely consistent with our theoretical story. In the US survey, as the response-length cutoff for defining compliance with the perspective-taking treatment increases, the effect estimates for respondents exposed to the China de-escalate treatment becomes larger, meaning those who complied with the treatment were more likely to prefer de-escalatory policies. Similarly, in the Chinese survey, as the response-length cutoff for defining compliance with the perspective-taking treatment increases, the effect estimate for respondents exposed to the US de-escalation treatment becomes larger, as is the case with the ratio estimator for respondents exposed to the US escalation treatment. Noncompliance thus makes the ATE results reported in the main text conservative estimates. Importantly, the estimates in [Tables](#)

3-4 assume one-sided noncompliance. It is also possible for there to be two-sided noncompliance, in that some participants in the control condition may already be engaged in perspective-taking regardless of treatment assignment; in a potential outcomes framework, these would be “always takers”. However, in this case the ratio estimator provides a lower bound on the effect, since non-compliance in the control group decreases the size of the compliance group (reflected in the denominator), increasing the estimated CACE, such that the results presented here remain conservative tests.

4.2 Balance tests

Figures 1 and 2 show that randomization was successful, in that respondents in both the Chinese and the US experiments are well-balanced across treatment conditions for a wide range of demographic characteristics.

Figure 1: Balance tests: China experiment

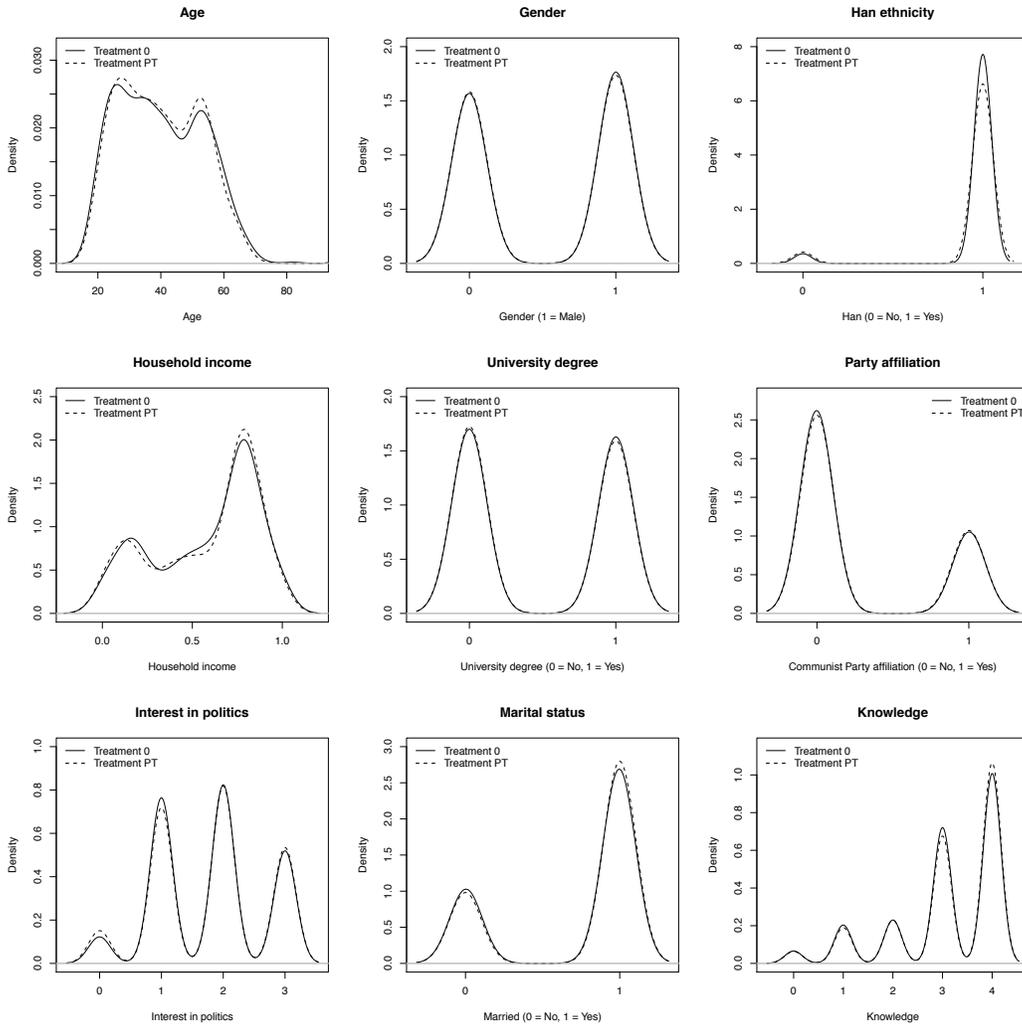
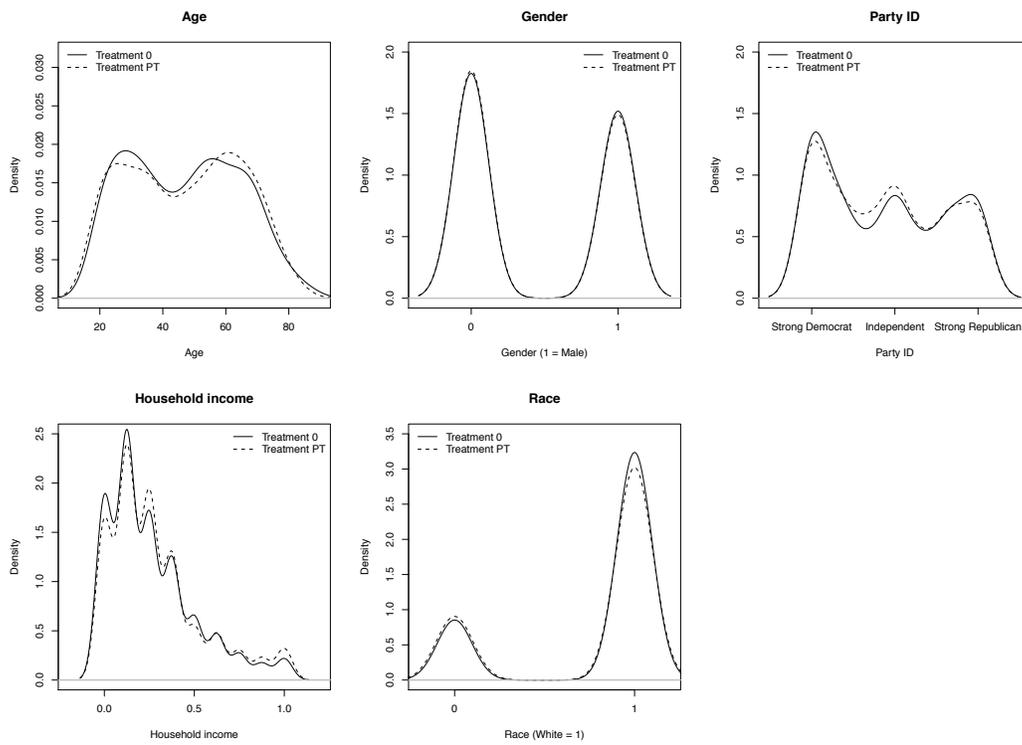


Figure 2: Balance tests: US experiment



4.3 Exploiting natural variation in empathy

The results from the paper only look at the effect of perspective-taking when it is experimentally induced, and we know that perspective-taking is a faculty that some respondents are naturally better at employing than others (Davis, 1983; Baron-Cohen and Wheelwright, 2004; Wakabayashi et al., 2006; Konrath, O'Brien and Hsing, 2011). How is this natural variation in empathy associated with escalation preferences? Does empathy have palliative effects when studied dispositionally, rather than induced experimentally? The top half of Table 5 presents results from a series of regression models in the US sample, and the bottom half for the China sample, both of which estimate the effect of dispositional levels of empathy controlling for the perspective-taking treatment and a variety of dispositional variables.

When participants are in the experimental condition where the other side de-escalates, we find that respondents' pre-dispositional levels of empathy have relatively little effect on their baseline policy preferences. Models 1-3 in Table 5 show that empathy has no main effect on policy preferences and no interaction effect with our perspective-taking treatment when the other side de-escalates, for both the United States and China. In contrast, we find that measures of military assertiveness and national chauvinism for both Americans and Chinese are consistently and predictably associated with preferences for escalation, and, as shown in previous analyses, experimentally inducing perspective-taking has a significant effect – inducing de-escalation among American respondents, and escalation among the Chinese.

When we shift our attention to respondents in the experimental condition where the other side escalates (models 4-6), we find stronger evidence of a relationship between dispositional empathy and foreign policy preferences, but not in the direction predicted by the palliative effects hypothesis. First, in the Chinese sample, models 4 and 5 of Table 5 show that more empathetic participants are *more* inclined to escalate. Second, dispositional empathy moderates the effect of the perspective-taking treatment among American respondents. In our previous analysis in the paper, we found that the perspective-taking treatment significantly affected escalation preferences in 3 of 4 subsamples of the experiment — the exception being American participants in the experimental condition where the other side escalates. The significant negative interaction term in model 6 of Table 5 suggests an interesting explanation for this effect. Substantively, the perspective-taking treatment has no significant effect on policy preferences among those respondents who are naturally predisposed to

Table 5: The dispositional empathy results offer little support for the palliative effects of empathy hypothesis (I)

Sample: US	China de-escalates			China escalates		
	(1)	(2)	(3)	(4)	(5)	(6)
Dispositional empathy	-0.201 (0.298)	-0.029 (0.276)	-0.124 (0.381)	-0.100 (0.313)	0.146 (0.277)	0.832** (0.386)
Perspective-taking	-0.268** (0.116)	-0.217** (0.108)	-0.366 (0.424)	-0.054 (0.121)	-0.020 (0.107)	1.022** (0.424)
Military assertiveness		2.754*** (0.280)	2.752*** (0.280)		3.350*** (0.269)	3.333*** (0.268)
National chauvinism		0.723*** (0.233)	0.726*** (0.233)		1.163*** (0.236)	1.150*** (0.235)
Education		0.167 (0.218)	0.165 (0.218)		0.253 (0.213)	0.233 (0.212)
Dispositional empathy × PT			0.201 (0.552)			-1.400** (0.551)
Constant	4.518*** (0.234)	2.360*** (0.297)	2.431*** (0.355)	4.835*** (0.249)	2.021*** (0.295)	1.533*** (0.351)
N	720	708	708	774	759	759
Adjusted R ²	0.005	0.159	0.158	-0.002	0.233	0.239
Sample: China	US de-escalates			US escalates		
	(1)	(2)	(3)	(4)	(5)	(6)
Dispositional empathy	0.422 (0.317)	0.067 (0.314)	0.439 (0.447)	0.660** (0.261)	0.448* (0.256)	0.137 (0.369)
Perspective-taking	0.248** (0.111)	0.236** (0.108)	0.742* (0.445)	-0.226** (0.096)	-0.176* (0.093)	-0.573 (0.352)
Military assertiveness		1.207*** (0.224)	1.198*** (0.225)		1.325*** (0.180)	1.334*** (0.180)
National chauvinism		0.073** (0.029)	0.072** (0.029)		0.065*** (0.023)	0.063*** (0.023)
Education		0.741** (0.366)	0.735** (0.366)		0.225 (0.336)	0.199 (0.337)
Dispositional empathy x PT			-0.726 (0.620)			0.589 (0.504)
Constant	4.865*** (0.237)	3.300*** (0.394)	3.054*** (0.446)	5.536*** (0.190)	4.152*** (0.311)	4.382*** (0.368)
N	754	753	753	728	728	728
Adjusted R ²	0.006	0.054	0.055	0.014	0.093	0.093

*p < .1; **p < .05; ***p < .01

engage in perspective-taking, but a strong positive effect among respondents who are dispositionally low in empathy. When the other side escalates, only those Americans least prone to engage in perspective-taking naturally need to be encouraged to do so. This is consistent with [Galinsky et al. \(2006\)](#), who find that power and perspective-taking are negatively correlated: when the US is in a position of relative strength in the region, respondents are less likely to naturally engage in perspective-taking unless encouraged to do so, whereas when the tables turn, only the least empathetic respondents need to be encouraged to engage in perspective-taking. Moreover, the sign of the effect is consistent with the reciprocity pattern reported in the US results above, since for those low-empathy respondents for whom the treatment has a significant effect, the effect is positive, mirroring Chinese behavior.

Table 6: The dispositional empathy results offer little support for the palliative effects of empathy hypothesis (II)

Sample: US	China de-escalates			China escalates		
	(1)	(2)	(3)	(4)	(5)	(6)
Dispositional empathy	-0.201 (0.298)	-0.031 (0.278)	-0.027 (0.383)	-0.100 (0.313)	0.140 (0.278)	0.827** (0.387)
Perspective-taking	-0.268** (0.116)	-0.210* (0.108)	-0.204 (0.426)	-0.054 (0.121)	-0.021 (0.107)	1.023** (0.424)
Military assertiveness		2.640*** (0.301)	2.640*** (0.301)		3.378*** (0.287)	3.360*** (0.286)
National chauvinism		0.644*** (0.234)	0.644*** (0.235)		1.166*** (0.237)	1.154*** (0.236)
Education		0.131 (0.218)	0.131 (0.218)		0.253 (0.214)	0.234 (0.214)
Male		0.235** (0.110)	0.235** (0.110)		-0.030 (0.109)	-0.035 (0.108)
Ideology		0.129 (0.220)	0.129 (0.220)		-0.046 (0.214)	-0.042 (0.214)
Dispositional empathy x PT			-0.008 (0.554)			-1.403** (0.552)
Constant	4.518*** (0.234)	2.316*** (0.306)	2.313*** (0.365)	4.835*** (0.249)	2.045*** (0.304)	1.557*** (0.359)
N	720	704	704	774	759	759
Adjusted R ²	0.005	0.159	0.158	-0.002	0.231	0.237
Sample: China	US de-escalates			US escalates		
	(1)	(2)	(3)	(4)	(5)	(6)
Dispositional empathy	0.422 (0.317)	0.016 (0.315)	0.417 (0.448)	0.660** (0.261)	0.494* (0.258)	0.208 (0.373)
Perspective-taking	0.248** (0.111)	0.214** (0.109)	0.758* (0.445)	-0.226** (0.096)	-0.183* (0.093)	-0.546 (0.355)
Military assertiveness		1.123*** (0.229)	1.112*** (0.230)		1.269*** (0.182)	1.279*** (0.182)
National chauvinism		0.062** (0.030)	0.061** (0.030)		0.072*** (0.023)	0.071*** (0.023)
Education		0.617* (0.369)	0.612* (0.369)		0.309 (0.343)	0.280 (0.344)
Male		0.116 (0.110)	0.117 (0.110)		0.241** (0.095)	0.235** (0.095)
Party Member		0.350*** (0.122)	0.349*** (0.122)		-0.119 (0.106)	-0.117 (0.106)
Dispositional empathy x PT			-0.781 (0.619)			0.539 (0.508)
Constant	4.865*** (0.237)	3.402*** (0.399)	3.136*** (0.452)	5.536*** (0.190)	3.952*** (0.318)	4.168*** (0.378)
N	754	747	747	728	720	720
Adjusted R ²	0.006	0.065	0.065	0.014	0.101	0.101

*p < .1; **p < .05; ***p < .01

Table 7: More empathetic respondents display larger attribution asymmetries (I)

	US sample			China sample		
	(1)	(2)	(3)	(4)	(5)	(6)
Dispositional empathy	0.599*	0.751**	1.145***	1.141***	0.995***	1.069**
	(0.328)	(0.318)	(0.441)	(0.347)	(0.351)	(0.502)
Perspective-taking	-0.048	-0.026	0.581	0.068	0.073	0.171
	(0.128)	(0.123)	(0.487)	(0.125)	(0.124)	(0.489)
Military assertiveness		3.382***	3.382***		0.662***	0.660***
		(0.315)	(0.315)		(0.248)	(0.249)
National chauvinism		-0.352	-0.361		0.129***	0.129***
		(0.270)	(0.270)		(0.032)	(0.032)
Education		0.809***	0.808***		-0.932**	-0.930**
		(0.247)	(0.247)		(0.433)	(0.433)
Dispositional empathy \times PT			-0.817			-0.143
			(0.634)			(0.691)
Constant	0.517**	-1.631***	-1.916***	2.070***	1.290***	1.237**
	(0.260)	(0.340)	(0.405)	(0.257)	(0.432)	(0.501)
N	1,494	1,467	1,467	1,480	1,479	1,479
Adjusted R ²	0.001	0.082	0.082	0.006	0.024	0.024

*p < .1; **p < .05; ***p < .01. See Table 8 for additional analysis.

Table 8: More empathetic respondents display larger attribution asymmetries (II)

	US Sample				China Sample						
	China de-escalates	(2)	China escalates	(3)	US de-escalates	(4)	US escalates	(5)	(6)	(7)	(8)
Dispositional empathy	0.102 (0.459)	0.223 (0.452)	1.059** (0.468)	1.251*** (0.450)	0.898* (0.495)	0.707 (0.496)	1.359*** (0.491)	0.898* (0.495)	0.707 (0.496)	1.359*** (0.491)	1.094** (0.506)
Perspective-taking	-0.081 (0.179)	-0.094 (0.175)	0.002 (0.181)	0.043 (0.173)	0.052 (0.173)	0.062 (0.171)	0.080 (0.181)	0.052 (0.173)	0.062 (0.171)	0.080 (0.181)	0.081 (0.183)
Military assertiveness		2.270*** (0.489)		3.683*** (0.465)		0.790** (0.362)					0.668* (0.358)
National chauvinism		-0.641* (0.380)		-0.098 (0.383)		0.205*** (0.046)					0.049 (0.046)
Education		0.800** (0.354)		0.895** (0.347)		-1.623*** (0.582)					-0.100 (0.673)
Male		0.123 (0.178)		-0.215 (0.176)		-0.015 (0.173)					-0.226 (0.186)
Ideology		0.726** (0.357)		0.668* (0.347)							
Party member						0.096 (0.192)					0.056 (0.208)
Constant	0.765** (0.360)	-0.993** (0.497)	0.277 (0.373)	-2.496*** (0.492)	2.257*** (0.371)	1.233* (0.629)	1.906*** (0.357)	2.257*** (0.371)	1.233* (0.629)	1.906*** (0.357)	1.435** (0.625)
N	720	704	774	759	753	746	727	753	746	727	719
Adjusted R ²	-0.002	0.051	0.004	0.116	0.002	0.044	0.008	0.002	0.044	0.008	0.008

*p < .1; **p < .05; ***p < .01

5 Exploring cross-national differences

There are three potential sets of reasons why we might see divergent patterns of results between the two countries. The first are *compositional differences* between the samples. If American samples differ from their Chinese counterparts in a variety of ways — their levels of nationalism, attitudes towards the use of force, and so on — and these dispositional traits affect how respondents react to the treatments, this could lead to divergent patterns of results between the two populations. The second are *relational differences* between the two countries. As [Boulding \(1959\)](#), [Herrmann and Fischerkeller \(1995\)](#) and others have noted, how you conceive of your country’s relationship with another actor — whether you see it as posing a threat or an opportunity, whether you see it as equal or inferior in capabilities, and so on — has important implications for the strategic scripts you propose. If there are asymmetries in the images respondents in each sample have of the other country, and these differences interact with the treatments, this could similarly lead to divergent outcomes between the two. Finally, the third are *situational differences* between the two countries — not reflecting divergent dispositional traits or belief systems between the two samples, but rather, the asymmetric stakes for the two countries themselves in a potential dispute in the South China Sea.

5.1 Compositional differences

We begin by testing for compositional differences. While there are a number of potential dispositional traits on which the two samples may differ, we identify four traits we consider to be particularly theoretically relevant, some of which come from the public opinion literature, and others of which hail from social psychology and behavioral economics. The first is attitudes towards the use of military force, or what the public opinion about foreign policy literature often calls militant assertiveness ([Herrmann, Tetlock and Visser, 1999](#); [Kertzer and Brutger, 2016](#)), the distinction between hawks and doves, who vary in their beliefs about the desirability and efficacy of the use of force. The second is national attachment, reflecting the extent to which respondents identify with their country ([Herrmann, Isernia and Segatti, 2009](#)); more nationalist individuals are not only more likely to feel like being a member of their national ingroup is an important part of their identity, but to have a starker perception of shared fate ([Brewer and Brown, 1998](#); [Herrmann, 2017](#)). Both national attachment and military assertiveness are worth investigating given widespread debates in foreign

policy circles about the nature and extent of Chinese nationalism and assertiveness (Johnston, 2013, 2017), and the demonstrated importance of each construct in the study of support for the use of force (Herrmann, Tetlock and Visser, 1999; Herrmann, Isernia and Segatti, 2009).

In the US sample, we measure attitudes towards the use of force with three items from the existing literature: two capturing beliefs about the efficacy of the use of force in general ("Going to war is unfortunate but sometimes the only solution to international problems", "The use of military force only makes problems worse"), and the third specifically implicating the role of one's own military ("The best way to ensure peace is through American military strength") (Kertzer and Renshon, 2014; Rathbun et al., 2016). Both are Likert items, scaled from "Strongly agree" to "Strongly disagree," and reverse-coded where applicable such that higher answers correspond to higher levels of hawkishness. In the Chinese sample, our military assertiveness measure consists of two of these items, one capturing general beliefs about the efficacy of the use of force ("Going to war is unfortunate but sometimes the only solution to international problems") and the other focusing specifically on the Chinese military ("The best way to ensure peace is through Chinese military strength"). In the analysis below, we therefore operationalize attitudes to the use of force both using construct-level comparisons (additive scales utilizing all of the military assertiveness items measured in each country) and item-level comparisons (focusing solely on items fielded in both samples).¹

We measure national attachment in the United States with two items from the national attachment scale of Herrmann, Isernia and Segatti (2009) ("When someone says something bad about the American people, how strongly do you feel it is as if they said something bad about you?" and "How much do you feel that what happens to America in general will be your fate?"); for our Chinese sample, we use the same question tapping national identification ("When someone says something bad about the Chinese people, how strongly do you feel it is as if they said something bad about you?"), while also using two standard measures of national pride commonly used in studies of Chinese nationalism ("Are you proud to be a Chinese citizen?", with responses ranging from "Not proud at all" to "Very proud", and "Is China superior to other countries?", with responses ranging from "Not superior" to "Superior by far"; see Johnston (2017) for similar items). As with the case of attitudes towards the use of force, then, in the analysis below, we study the effects of nationalism

¹As one expects, reliability estimates are higher for the three-item scale than the two-item scale: In the US sample: $\alpha = 0.69$; in the Chinese sample: $\alpha = 0.52$.

both using construct-level and item-level comparisons.².

The third trait comes not from the public opinion literature, but behavioral research on strategic thinking. Behavioral economists often measure strategic thinking through “beauty contest games” (Camerer, Ho and Chong, 2004), based on a discussion in Keynes (1936) comparing the stock market to a beauty contest where players win by accurately guessing which contestants other players will deem to be the most beautiful. In one common version of the game (Coricelli and Nagel, 2009; LeVeck et al., 2014), players are asked to guess a number between 0 to 100, in which the winner will be the player whose guess is the closest to some multiplier (e.g. two-thirds) of the average guess of the group. The guesses are then scored as a measure of *k-level reasoning*: level-0 reasoners guess randomly between 0-100, while level-1 reasoners assume the other players will guess randomly such that the best response is 33 ($50 \times \frac{2}{3}$), level-2 reasoners assume other players are capable of $K - 1$ -level reasoning such that the best response is 22 ($50 \times \frac{2}{3}^2$), and so on. In this sense, k-level reasoning is often used as a measure of “strategic skill” in that level-two reasoners are more strategic than level-one reasoners, and so on (Hafner-Burton et al., 2014). Yet it also measures strategic players’ implicit beliefs about the other players. While the equilibrium of the game is 0, most players in the real world don’t approximate *homo economicus* and iteratively eliminate dominated strategies all the way; winning the game thus requires respondents to engage in strategic empathy and attempt to gauge how strategic the other players are. Mindful of these dual meanings, we thus adapted the measure of k-level reasoning used by LeVeck et al. (2014), but with a cross-national twist. For half the respondents in each country, we presented the game as follows:

Pretend you’re playing a game against other players, where every participant has to think of a whole number from 0 to 100. The winner of the game is the player whose guess is 60% of the average guess of the group. (For example, if the average guess of the group is 90, the winner is the participant in the group who guessed the number closest to 54).

What number would you pick?

For the other (random) half of respondents in each country, we repeated the same prompt, but modified the first sentence to begin with “Pretend you’re playing a game against players from China” for the US respondents, or “players from the United States” for the Chinese respondents. By randomizing the identity of the opponents, we can test whether respondents’ k-level reasoning

² $\alpha = 0.5$ in the US, $\alpha = 0.67$ in China

scores differ when playing against opponents from the opposing country. Regardless of treatment condition, we analyze the results the same way, following [LeVeck et al. \(2014\)](#) by differentiating between level- $K = 0$, $K = 1$, $K = 2$, and $K = 3$ (or greater) respondents. As it turns out, we find that the identity treatment does not affect respondents' k-level scores in either country ($p < 0.624$ for the US sample, $p < 0.858$ for the Chinese sample, each from a Wilcoxon-rank sum test), so in the analysis below we combine k-level scores across identity treatment conditions.

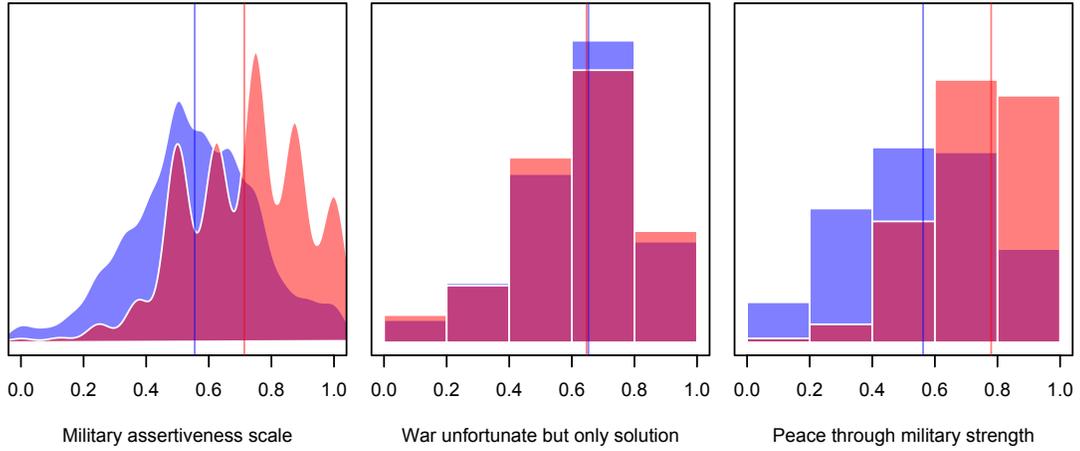
Finally, although perspective-taking is studied situationally through random assignment in the main text, we can also study it as an individual-difference, using the dispositional measure from [Davis \(1980\)](#) described elsewhere in the appendix. As with the other respondent-level variables studied above, our empirical strategy here has two steps. First, on which traits do the two samples significantly differ from one another? Second, do respondents with varying levels of these traits respond differently to the perspective-taking treatment?

Figure 3 presents the results of the first stage. Panel a shows that the Chinese sample is significantly more hawkish than the American counterpart ($t = -21.9$, $p < 0.000$); item-level comparisons suggest the difference is not attributable to beliefs about the efficacy of force in general, but because Chinese respondents are significantly more likely to agree that the best way to ensure peace is through Chinese military strength than American respondents are about American military strength ($W = 622460$, $p < 0.000$). Panel b shows that although the difference is not as stark as that with respect to attitudes towards the use of force, the Chinese sample is also significantly more nationalist ($t = -3.83$, $p < 0.001$) than the American sample; the item-level comparison similarly finds that Chinese respondents are significantly more likely to feel that when someone says something bad about your country, it is as if they said something bad about you, although the Chinese-language survey had a four-point rather than five-point set of response options for this item, such these differences could be due to measurement properties rather than anything substantive. Panel c shows that American respondents displayed slightly higher k-level reasoning than Chinese respondents ($W = 1127300$, $p < 0.000$), though the differences are small; 84.0% of American respondents were 0-level reasoners, while 89.7% of Chinese respondents were 0-level reasoners. Finally, panel d suggests that when we turn to a dispositional measure of strategic empathy, our American respondents are higher in perspective-taking than our Chinese respondents ($t = 8.76$, $p < 0.000$).

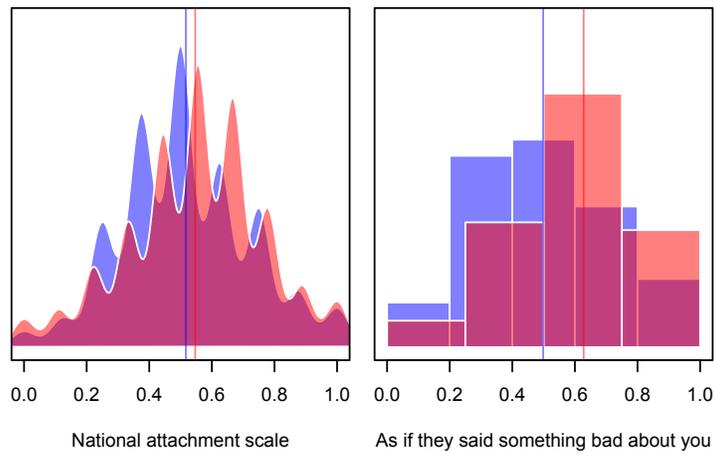
While varying in size, then, we find evidence of compositional differences between the two samples for all four traits. To determine whether they can account for the divergent pattern of

Figure 3: Compositional differences across samples

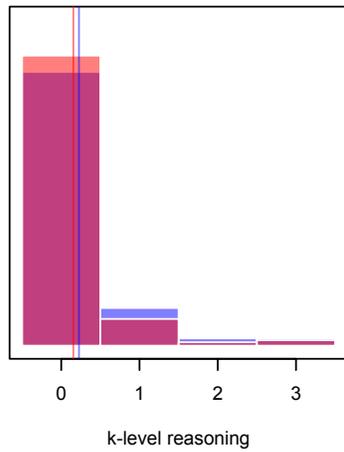
(a) Militant assertiveness



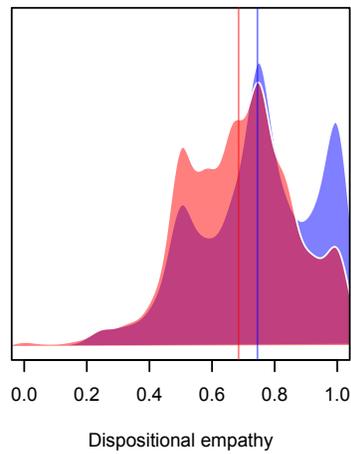
(b) National attachment



(c) k-level reasoning

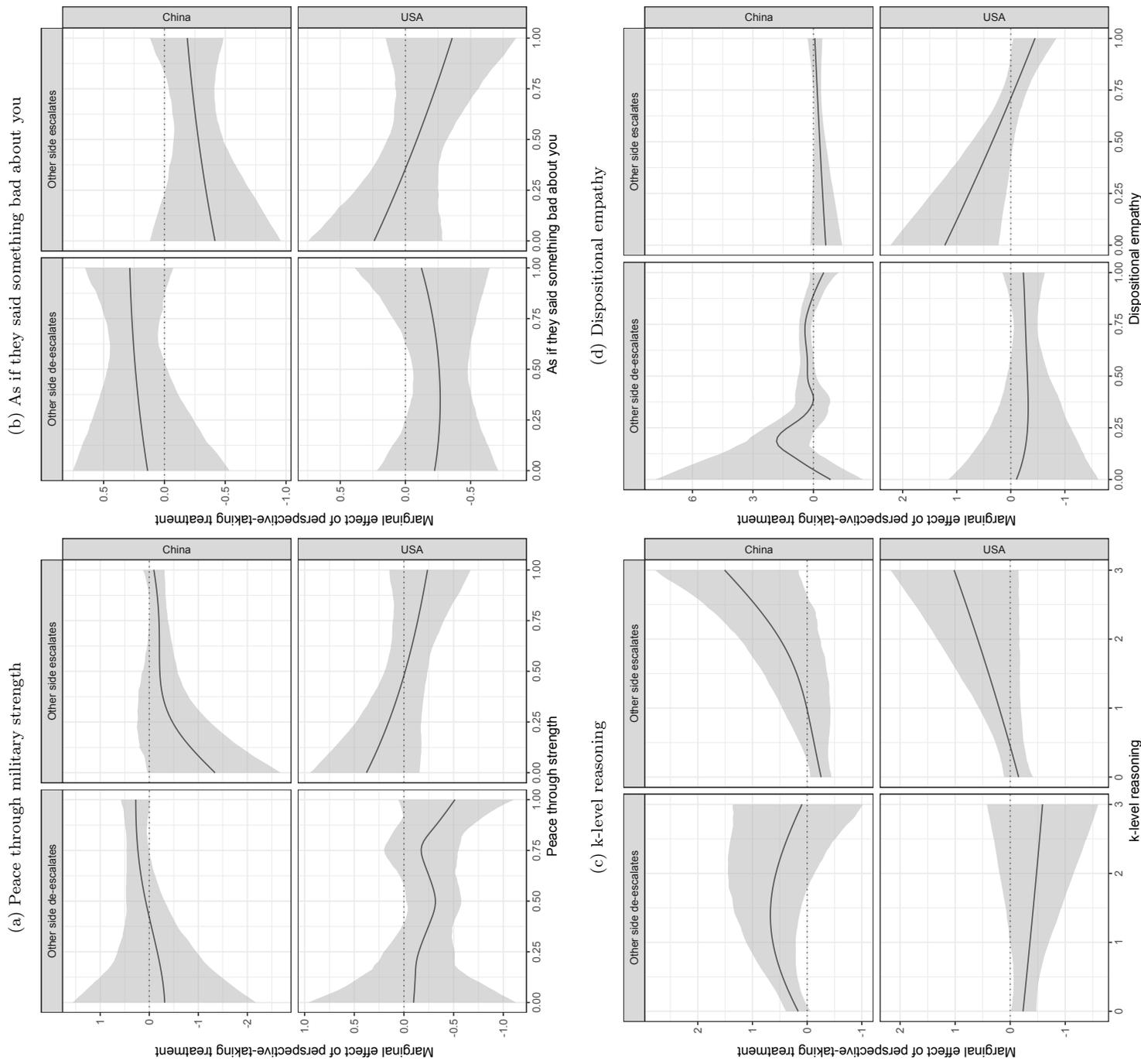


(d) Dispositional empathy



Note: US sample in blue, Chinese sample in red. The mean of each distribution is represented by a vertical line.

Figure 4: Heterogeneous treatment effects across dispositional variables



Note: each panel displays the marginal effect of the perspective-taking treatment across a different dispositional variable, estimated using a kernel estimation procedure with 5-fold least-squares cross-validation using the `interflex` package in R to avoid making assumptions about the functional form of the interactive effect.

results, we then estimate the marginal effect of the perspective-taking treatment conditional on each trait in question; to avoid imposing a linear functional form, we use a kernel estimation procedure introduced by [Hainmueller, Mummolo and Xu \(2018\)](#), which uses 5-fold least-squares cross-validation to flexibly estimate the functional form of the interaction; in order to make consistent a comparison as possible between each sample, we use the item-level comparisons for militant assertiveness and national attachment.

Figure 4 presents the marginal effects. Panel (a) shows the results for the military assertiveness item about producing peace through military strength, which the Chinese sample scored significantly higher in than the American sample. Yet the pattern of marginal effects cannot explain the divergent results: it is not the case that more dovish Chinese respondents (those in the left half of each panel in the top row) display perspective-taking-effects that more closely resemble those in the US sample. For example, it looks as if peace through strength moderates the effect of perspective-taking in the Chinese sample when the US escalates, but the marginal effect appears negative, rather than positive as with American respondents. Panel (b) shows the results for the national attachment item about feeling that when someone says something bad about your country, it is if they said something bad about you (an item the Chinese sample scores higher in than the American sample), but the marginal effects appear either relatively flat, or not statistically significant. Panel (c) turns to k-level reasoning. Here we see significant interaction effects – for example, in the Chinese sample when US escalates, perspective-taking decreases the likelihood of escalation among the lowest-level reasoners, while increasing it among the highest-level ones. But, this cannot explain the patterns of results we see in the US. Finally, panel (d) looks at dispositional empathy. Here, we see a strong linear interaction effect for the US sample when the other side escalates — corresponding with the results presented elsewhere in the Appendix — but it is not the case that the less strategically empathetic American respondents in the left-hand half of that panel behave more like their Chinese counterparts, given that the effect of perspective-taking amongst Chinese respondents in a similar condition is negative. Thus, although there are compositional differences between the samples, they cannot explain the divergent results we see here.

5.2 Relational differences

We then turn to testing for relational differences between the two samples: asymmetries in the images citizens in each country have about the other. The central insight of image theory in IR is

that, as is the case with stereotypes more generally, the perceptions actors have of other countries in international politics tend to be structured along a small number of dimensions (Herrmann, 2013; Fiske et al., 2002). These “combinatorial constructs” matter, because the images actors have of others correspond with specific strategic scripts (Herrmann and Fischerkeller, 1995; Castano, Bonacossa and Gries, 2016). Thus, differences in how Chinese and American respondents conceive of the relationship between the two countries could lead to divergent policy preferences as well.

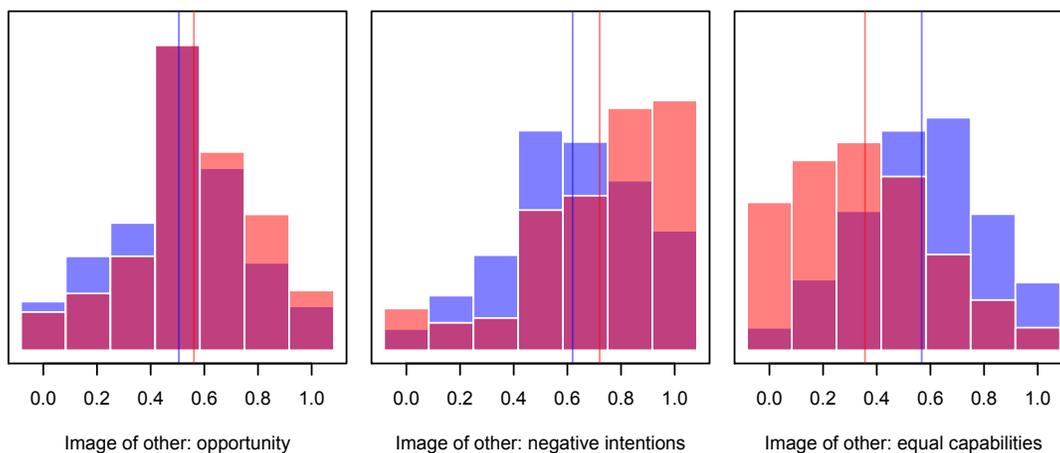
In the most prominent version of image theory in IR, actors’ perceptions of other countries are understood as being structured along three dimensions: the degree of perceived threat or opportunity, relative capabilities, and cultural status (Herrmann and Fischerkeller, 1995). Given controversies about the role of the cultural status dimension (stereotype content models, for example, are typically two-dimensional - see Herrmann, 2013), we focus on two of these here. First, we asked our respondents two items getting at perceptions of threat or opportunity: is the other actor a crucial partner for cooperation? Or, a dissatisfied power with expansionist ambitions? To this end, respondents were asked to indicate the extent to which they agree with the statements “[China/the United States] is China’s most important partner”, and “[China/the United States] wants to dominate the world.” Second, we asked our respondent an item soliciting their perception of the balance of power between the two countries, asking them the extent to which they agree with the proposition that “With regard to global power, China has closed the gap with the United States”. All three Likert responses are scaled such that higher values indicate stronger agreement.

In addition to these image theory measures, we also use our attribution measures (described in detail in the main text) as a different way of operationalizing how respondents think about the relationship between the two countries, focusing in particular on how offensively motivated respondents perceive escalation by the other side to be, and the asymmetry in attribution offered between the other side’s escalation and your own side’s escalation. These items less strictly capture relational differences than the image theory items, since they focus on how respondents understand the relationship in the specific context of escalation in the South China Sea, rather than as a more general construct. Nonetheless, they are still valuable in as much divergences in how the two sides interpret each other’s behavior in the region provides insight into their divergent responses to the perspective-taking treatment in the main experiment.

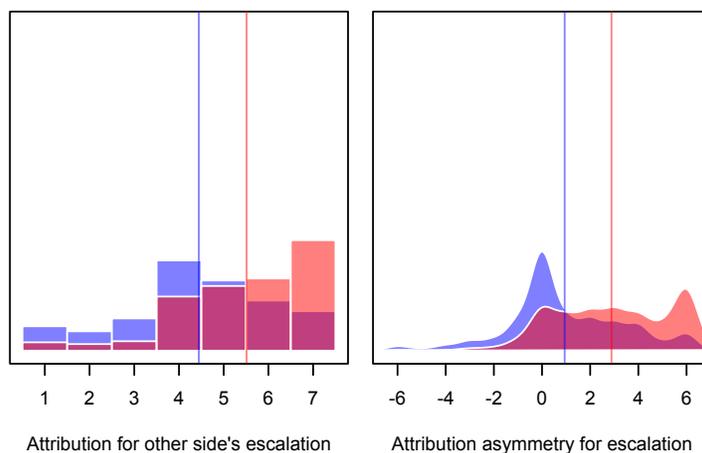
As before, our empirical strategy here has two steps. First, is there an asymmetry in how respondents from the two samples perceive of their countries’ relationship? Second, does variation

Figure 5: Relational differences across samples

(a) Image of the other side



(b) Attributions for escalation

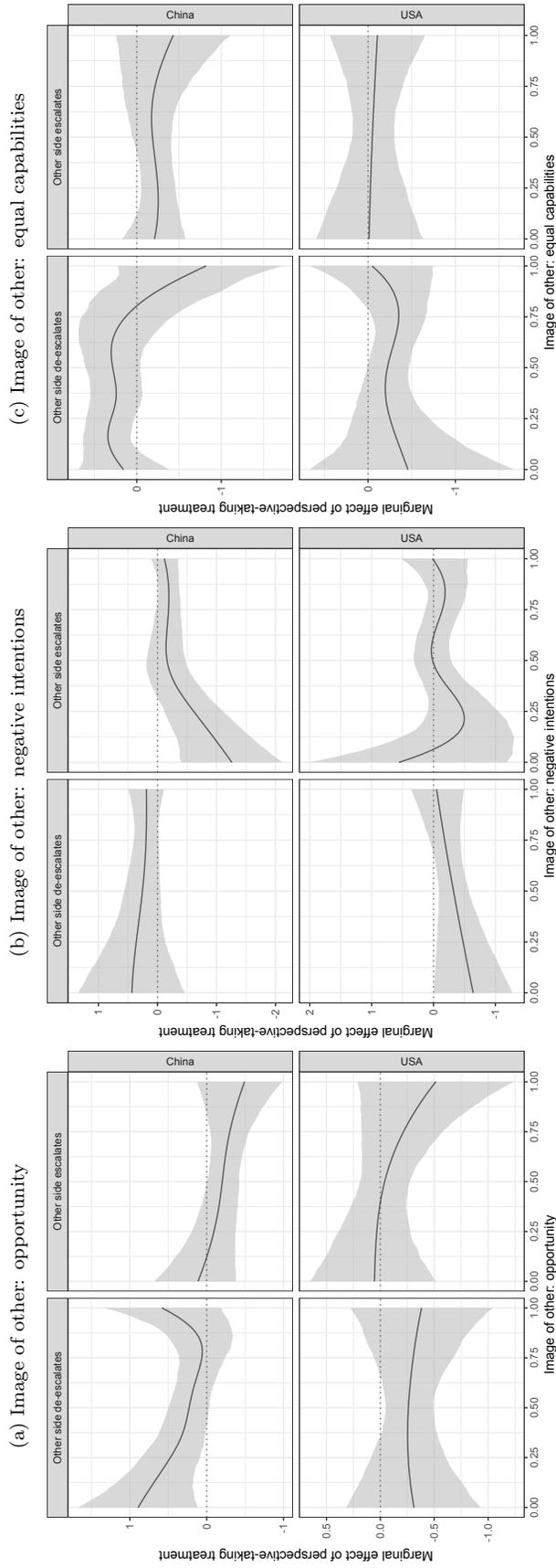


Note: US sample in blue, Chinese sample in red. The mean of each distribution is represented by a vertical line.

in the images perceived by each side shape how they respond to the perspective-taking treatment? Figure 5 presents the results of this first step, with panel a depicting the distributions of image scores across two the samples. The first panel shows that both populations perceive each other as a relatively important partner, but that there are asymmetries in perceptions of interdependence: Chinese respondents are significantly more likely to see the US as China’s most important partner than vice versa ($W = 958910, p < 0.000$). The second panel displays an even more striking gap in the extent to which respondents perceive the other side as being driven by negative, expansionist ambitions, with Chinese respondents being significantly more likely to perceive the United States as harboring threatening intentions than Americans do so with China ($W = 827560, p < 0.000$). A similarly stark difference is shown in the third panel: American respondents are much more likely to perceive China as having caught up to the United States in terms of global power than Chinese respondents do ($W = 1604200, p < 0.000$). Thus, respondents in each country not only perceive each other very differently, but also perceive the nature of their relationship in different ways.

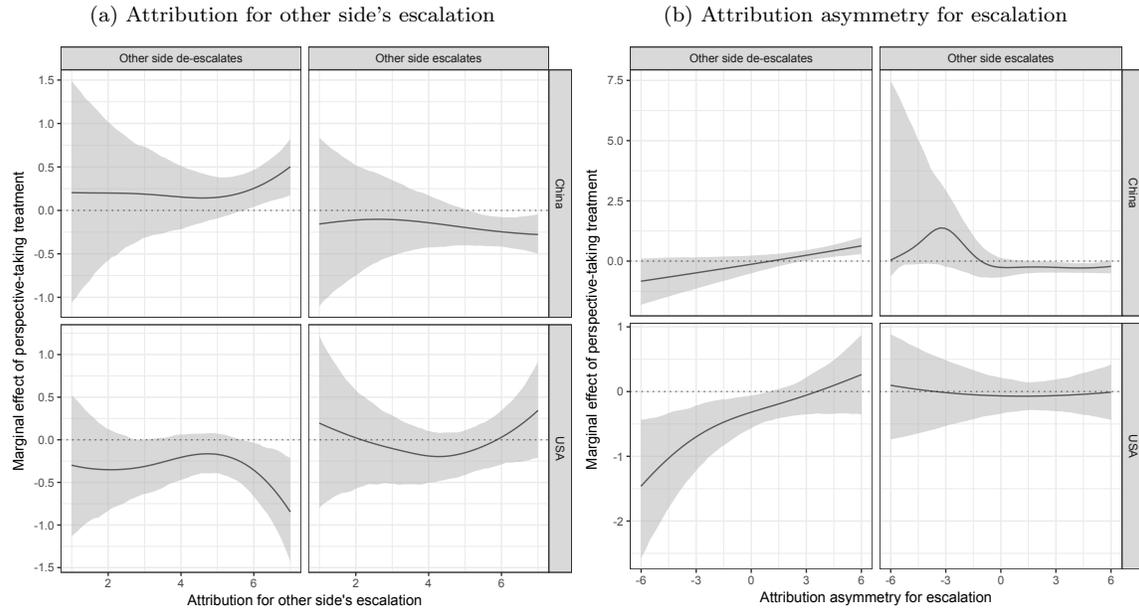
Panel b turns to the attribution measures. Here, the items are scaled such that higher values indicates a more hostile interpretation of the other side’s motives, in either absolute (the first panel) or relative terms (the second panel). As is evident in the attribution analysis in the main text, and presented graphically in the first panel here, there are significant differences in the interpretation Chinese and American respondents have for each other’s escalation in the South China Sea: American respondents are less likely to perceive Chinese escalation as offensively-motivated than Chinese respondents are with American escalation ($W = 689340, p < 0.000$): for American respondents, the modal response is the scale midpoint, while for Chinese respondents, the modal response is the scale maximum. As we suggest in the next section, these divergent attributions are likely due to the nature of the dispute, and its geographic proximity to China. The second panel shifts from absolute to relative attributions, depicting the distribution of respondent-level differences in attributions for the other side escalating versus their own side escalating. Positive values indicate an attribution asymmetry in which the other side’s behavior is seen as more offensively-motivated than when the same action is carried out by the other side, while negative values indicate an attribution asymmetry in which the other side’s behavior is seen as more defensively-motivated than when the action is carried out by your own side. As noted in the main text, Chinese respondents display a significantly larger attribution asymmetry ($t = -21.76, p < 0.000$): compared to American respondents, they perceive the other side’s behavior as much more offensively-motivated than they perceive their own.

Figure 6: Heterogeneous treatment effects across image variables



Note: each panel displays the marginal effect of the perspective-taking treatment across a different relational variable, estimated using a kernel estimation procedure with 5-fold least-squares cross-validation using the `interflex` package in R to avoid making assumptions about the functional form of the interactive effect.

Figure 7: Heterogeneous treatment effects across attribution variables



Note: each panel displays the marginal effect of the perspective-taking treatment across a different relational variable, estimated using a kernel estimation procedure with 5-fold least-squares cross-validation using the `interflex` package in R to avoid making assumptions about the functional form of the interactive effect.

Thus, there are clear asymmetries in how respondents from each country perceive one another, and the nature of their relationship. The question is whether these asymmetries can help explain the asymmetrical responses our respondents display to the perspective-taking treatment. As before, we test this possibility by estimating the marginal effect of the perspective-taking treatment conditional on each relational variable, using 5-fold least-squares cross-validation to flexibly estimate the functional form of the interaction.

Figure 6 presents the marginal effects for the image variables. Panel a) suggests relatively little evidence of a significant interaction effect (though it appears that Chinese respondents who perceive the US as a more important partner are more likely to de-escalate in the face of American escalation when presented with the perspective-taking treatment), and Chinese respondents who perceive the US to be a less important partner do not respond to the treatment more similarly to American respondents. The second panel shows that Chinese respondents who perceive US intentions as less threatening are more likely to de-escalate in the face of US escalation; panel c) finds no evidence of significant interaction effects regardless of how respondents assess the balance of capabilities. In both panels, then, there is little evidence that heterogeneous images account for the pattern of results we

see across the two countries.

Figure 7 turns to the marginal effects for the attribution variables. While panel a) suggests relatively little evidence of heterogeneous treatment effects, panel b) displays some more striking results. US respondents who perceive positive attribution asymmetries — that is, who perceive Chinese escalation as relatively more offensively-motivated — are less likely to reciprocate Chinese de-escalation, making their behavior more like the Chinese respondents. Similarly, Chinese respondents who display lower attribution asymmetry (making them more similar to US respondents) no longer display a significant perspective taking effect (another pattern more similar to US respondents), although caution must be taken in overinterpreting the result given the skewed distribution of attribution asymmetry among Chinese respondents. Nonetheless, compared to all of the compositional and all other relational tests, there is some evidence that variation in the relative degree of offensive motivations attributed to the other side for escalating contributes to some of the divergent responses we see in the experiment. Since, as noted above, these attribution measures focus on escalation attributions in the context of the South China Sea specifically, this suggests there is something about the nature of the situation itself, which we turn to next.

5.3 Situational differences

We thus turn to potential situational differences. Chinese and U.S. respondents may respond to perspective-taking differently because they each view and value the South China Sea differently. On the one hand, among Americans, the South China Sea is a distant area that is not of great importance to the United States. Importantly, Americans do not have any perceived ownership over the South China Sea. This strategic view was emphasized by one of the U.S. respondents, who noted “It isn’t our area to dictate to the world who should or shouldn’t have jurisdiction there.” On the other hand, among the Chinese, it is popularly perceived that China has a rightful claim over the South China Sea – a perception that is constantly reinforced by the government-controlled media. As our Chinese respondents reiterated: “China takes a strong stand and attaches great importance to the South China Sea issue” and “China’s claim to the South China Sea is consistent and China will uphold sovereignty by all means.” Another put it vividly: “China and the United States are located on two sides of the planet. On what grounds does the United States have to come all the way over and intervene in our territory?”

The perceived stakes in the dispute are thus notably different for Americans and the Chinese.

Since Americans do not view the South China Sea as part of the United States, and there is no interest to take the territory even if the United States were capable to do so, our American respondents viewed the strategic goal primarily as maintaining the status quo. As one respondent put it, "I would not want the U.S. to appear to be an aggressor, however the U.S. should send a signal that we are watching and will not stand by if China makes a very aggressive move." Another was more direct, simply commenting "Keep the status quo." If the U.S. public responds based on this worldview, and we expect perspective-taking to strengthen reliance on prior beliefs, then in line with deterrence and status quo maintenance, we should expect perspective-taking to strengthen beliefs that China is cooperative (or not) in terms of maintaining the status quo, depending on China's action in the experimental scenario. In other words, if Americans see China de-escalating, they would infer China is willing to cooperate and would therefore support U.S. de-escalation to maintain the status quo; if Americans see China escalating, they would infer China is uncooperative and would therefore support escalation to deter China from changing the status quo.

To the Chinese, the South China Sea is part of China's territory, and it is a rightful ownership claim that should be materialized.³ However, the Chinese also hope to avoid a costly conflict with the United States.⁴ The Chinese public is generally well aware that the United States is militarily stronger, and is rightly concerned about U.S. potential reactions if China attempts to revise the status quo.⁵ If the Chinese public responds based on this worldview, we should expect perspective-taking to strengthen beliefs that the United States is resolved (or not) in defending the status quo, depending on the U.S. action in the experimental scenario. U.S. escalation would be seen as a sign of strength, deterring the Chinese public from supporting escalation and risking a direct conflict with the United States, whereas de-escalation would be seen by the Chinese as a sign of weakness, which allows them to defend their rightful claim to the islands. As a Chinese respondent wrote: "The fact that the United States decided to withdraw from the South China Sea issue shows that our country

³See, e.g. Gao Jun, "Ting, Nanhai zhi ming" [Listen, the Name of the South China Sea], *Renmin Ribao* [People's Daily], June 26, 2014; Zhong Sheng, "Zhongguo dui nanhai zhudao de zhuquan burong fouding" [China's sovereignty rights over islands in the South China Sea cannot be denied], *Renmin Ribao* [People's Daily], December 15, 2015; Hu Xize and Zhang Mengxu. "Lishixing quanli liying youxian kaolu" [Historical rights should be prioritized in the consideration], *Renmin Ribao* [People's Daily], July 23, 2016. See also Ministry of Foreign Affairs, "Statement of the Government of the People's Republic of China on China's Territorial Sovereignty and Maritime Rights and Interests in the South China Sea," July 12, 2016.

⁴See, e.g., "Zhengshi dashi, buwang chaju – ting haiwai zhuanjia tan 'zhongguo shiji'" [Facing the trend, not forgetting the gap – Listening to foreign experts on the "China Century"], *Xinhua Net*, January 1, 2015; Zheng Zhenqing, "Wei shijie tigong gengduo guoji gonggong chanpin" [Providing more international public good for the world], *Renmin Ribao* [People's Daily], January 7, 2018.

⁵See, e.g., "Chinese Public Sees More Powerful Role in World, Names U.S. as Top Threat," Pew Research Center, October 2016.

has really become powerful. The United States is no longer the hegemonic power that can intervene with any country and any affair in the world, by whatever measures they want to take! I am really proud to be born as a Chinese!" Other respondents expressed similar views, pointing out that the United States "is not as powerful as before in dealing with China"; that the United States "pays much more attention to China's military strength. China's comprehensive national power keeps rising, and its international status is becoming more and more important"; and that the United States "is a country that is outside the region, it should not get involved in here at the first place. Therefore, its withdrawal is normal and wise." Another respondent put it colorfully: "China is no longer the "Sleeping Lion", the "Dragon of the East" has begun to take off! The Americans . . . have to acknowledge the undeniable fact that they are no longer the hegemon who can claim they are doing the "right thing" even if they are wrong."

These survey responses corroborate with our interviews with native Chinese observers, one of whom highlighted that the traditional Chinese military/war strategy is "敌退我进; 敌进我退", or "advance when the enemy retreats; retreat when the enemy advances." This pragmatic principle of taking actions based on estimation of the relative strengths of the two conflicting sides is reflected in the writings of the Chinese military strategist, Sunzi. Thus, "the Chinese tend to see U.S. de-escalation as a sign of its weakness; and U.S. escalation, a sign of strength and confidence." Another observer reiterated that the Chinese public perspective on the conflict is influenced by "the official interpretation of the history since 1840" and "typical slogans" that include "落后就要挨打" (we are bullied if we are weak) and "弱肉强食" (the weak are the prey of the strong). "The majority still regards the United States as the strongest power. Thus, the United States can do whatever it wants . . . without constraint." At the same time, "many people do not see the United States as a legitimate player in the game: we are defending the legitimate sovereignty rights over the islands, so we are purely defensive and not a threat to the region; the United States does not have territorial disputes in the [South China Sea], so it has no intrinsic interests; we will not invade the Philippines so the argument about defending U.S. ally is nonsense . . . the only reason for the United States to stay in the region and the dispute, in particular, is to contain China." In the end, "China's stake is high in the South China Sea, so it will incentivize [the Chinese] to pursue a decisive victory."

References

- Angrist, Joshua D, Guido W Imbens and Donald B Rubin. 1996. "Identification of causal effects using instrumental variables." *Journal of the American statistical Association* 91(434):444–455.
- Baron-Cohen, Simon and Sally Wheelwright. 2004. "The empathy quotient: an investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences." *Journal of autism and developmental disorders* 34(2):163–175.
- Boulding, Kenneth E. 1959. "National images and international systems." *Journal of Conflict Resolution* 3(2):120–131.
- Brewer, Marilynn B. and Rupert J. Brown. 1998. Intergroup Relations. In *The Handbook of Social Psychology*, ed. Daniel T. Gilbert, Susan T. Fiske and Gardner Lindzey. New York: McGraw-Hill pp. 554–594.
- Camerer, Colin F., Teck-Hua Ho and Juin-Kuan Chong. 2004. "A Cognitive Hierarchy Model of Games." *The Quarterly Journal of Economics* 119(3):861–898.
- Castano, Emanuele, Alain Bonacossa and Peter Gries. 2016. "National Images as Integrated Schemas: Subliminal Primes of Image Attributes Shape Foreign Policy Preferences." *Political Psychology* 37(3):351–366.
- Chang, Linchiat and Jon A. Krosnick. 2010. "Comparing Oral Interviewing with Self-Administered Computerized Questionnaires: An Experiment." *Public Opinion Quarterly* 74(1):154–167.
- Coricelli, Giorgio and Rosemarie Nagel. 2009. "Neural correlates of depth of strategic reasoning in medial prefrontal cortex." *Proceedings of the National Academy of Sciences* 106(23):9163–9168.
- Davis, Mark H. 1980. "A multidimensional approach to individual differences in empathy." *Catalog of Selected Documents in Psychology* 10:85.
- Davis, Mark H. 1983. "Measuring individual differences in empathy: Evidence for a multidimensional approach." *Journal of personality and social psychology* 44(1):113–126.
- Fiske, Susan T., Amy J. C. Cuddy, Peter Glick and Jun Xu. 2002. "A Model of (Often Mixed) Stereotype Content: Competence and Warmth Respectively Follow From Perceived Status and Competition." *Journal of Personality and Social Psychology* 82(6):878–902.
- Galinsky, Adam D., Joe C. Magee, M. Ena Inesi and Deborah H Gruenfeld. 2006. "Power and Perspectives Not Taken." *Psychological Science* 17(12):1068–1074.
- Gerber, Alan S. and Donald P. Green. 2012. *Field Experiments: Design, Analysis, and Interpretation*.

- New York, NY: W.W. Norton.
- Hafner-Burton, Emilie M, Brad L LeVeck, David G Victor and James H Fowler. 2014. "Decision maker preferences for international legal cooperation." *International Organization* 68(04):845–876.
- Hainmueller, Jens, Jonathan Mummolo and Yiqing Xu. 2018. "How much should we trust estimates from multiplicative interaction models? Simple tools to improve empirical practice." *Political Analysis* Forthcoming.
- Herrmann, Richard K. 2013. Perceptions and Image Theory in International Relations. In *Oxford Handbook of Political Psychology*, ed. Leonie Huddy, David O. Sears and Jack S. Levy. 2nd ed. ed. Oxford University Press pp. 334–363.
- Herrmann, Richard K. 2017. "How Attachments to the Nation Shape Beliefs About the World: A Theory of Motivated Reasoning." *International Organization* 71(S1):S61–S84.
- Herrmann, Richard K. and Michael P. Fischerkeller. 1995. "Beyond the Enemy Image and Spiral Model: Cognitive-Strategic Research After the Cold War." *International Organization* 49(3):415–50.
- 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):553–573.
- Herrmann, Richard K., Pierangelo Isernia and Paolo Segatti. 2009. "Attachment to the Nation and International Relations: Probing the Dimensions of Identity and their Relationship to War and Peace." *Political Psychology* 30(5):721–754.
- Johnston, Alastair Iain. 2013. "How New and Assertive Is China's New Assertiveness?" *International Security* 37(4):7–48.
- Johnston, Alastair Iain. 2017. "Is Chinese Nationalism Rising: Evidence from Beijing." *International Security* 41(3):7–43.
- Kertzer, Joshua D. and Jonathan Renshon. 2014. "Putting Things in Perspective: Perspective-Taking in Experimental Political Science." Working paper.
- 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):234–249.
- Keynes, John Maynard. 1936. *The General Theory of Interest, Employment and Money*. Macmillan.
- Konrath, Sara H., Edward H. O'Brien and Courtney Hsing. 2011. "Changes in Dispositional Empathy in American College Students Over Time: A Meta-Analysis." *Personality and Social Psychology Review* 15(2):180–198.

- LeVeck, Brad L., D. Alex Hughes, James H. Fowler, Emilie Hafner-Burton and David G. Victor. 2014. "The role of self-interest in elite bargaining." *Proceedings of the National Academy of Sciences* 111(52):18536–18541.
- Rathbun, Brian C., Joshua D. Kertzer, Jason Reifler, Paul Goren and Thomas J. Scotto. 2016. "Taking Foreign Policy Personally: Personal Values and Foreign Policy Attitudes." *International Studies Quarterly* 60(1):124–137.
- Wakabayashi, Akio, Simon Baron-Cohen, Sally Wheelwright, Nigel Goldenfeld, Joe Delaney, Debra Fine, Richard Smith and Leonora Weil. 2006. "Development of short forms of the Empathy Quotient (EQ-Short) and the Systemizing Quotient (SQ-Short)." *Personality and Individual Differences* 41(5):929 – 940.