Identity and the Social Construction of Reputation in World Politics

Jonathan Renshon *   Ryan Powers†

April 19, 2024

Abstract

Scholarship on reputations in IR has thus far left out one of the most striking features of human psychology: identity. Categorizing others as either "us" or "them" is an automatic and pervasive process that has significant implications for how reputations are generated and maintained. We provide a theoretical framework—based on social identity theory—to explain how identity dynamics affect both how we perceive other actors' "type" and estimate their likely behavior; in short, their reputations. Empirically, we provide two contributions. First, we field a descriptive survey on one public and two elite samples to learn about the content and boundaries of international identity and aid in the design of our experiment. Second, we field a pilot survey experiment that tests implications of our identity-based theory of reputations. We find that ingroup membership confers significant benefits to (1) reputations of all types (financial, signaling, humanitarian and resolve) and (2) observers' willingness to engage with the state, as well as (3) preliminary evidence that target behavior is differentially integrated into reputations for ingroup members compared to outgroups.

---

*Professor, Department of Political Science, University of Wisconsin-Madison. ✉: renshon@wisc.edu. 🖥: http://jonathanrenshon.com. This is one of several joint articles by the authors; the ordering of names reflects a principle of rotation.

†Assistant Professor, Department of International Affairs, School of Public and International Affairs, University of Georgia.. ✉: ryan.powers@uga.edu 🖥: https://ryanpowers.net.

The study of reputations is an incredibly rich, productive area in IR and political science more generally, and dates back to at least Schelling's (1966, 124) admonition that a country's "image"—their "reputation for action"—is "one of the few things worth fighting over." We've learned that reputations for resolve lower the odds that a state will be challenged in the future (Weisiger and Yarhi-Milo, 2015), that concern for reputation among leaders varies systematically (Yarhi-Milo, 2018) and can lead to more frequent and longer violent conflict (Dafoe and Caughey, 2016), and that these dynamics affect behavior across a variety of domains, including international economic cooperation (Chen, Pevehouse and Powers, 2023), civil conflict (Walter, 2009) and humanitarian interventions (Löwenheim, 2003). Debates about reputation have crossed rationalist and psychological approaches (Mercer, 1997, 2010) and generated several recent reviews that highlight much more progress than we can survey in this brief introduction (Dafoe, Renshon and Huth, 2014; Jervis, Yarhi-Milo and Casler, 2021).

While other aspects of reputation have been vigorously debated, there is a near-universal consensus about where reputations come from: they are formed and updated based on observed behavior. Yet, a number of historical cases suggest that this process is fraught with error. Russia's mistaken beliefs about Japanese resolve and "character" in the lead-up to their 1905 war is notorious (Streich and Levy, 2016), but Western observers, presumably incentivized to view events more accurately, fared no better (Mercer, 2023). During the Cold War, key leaders—notoriously John Foster Dulles—failed to update their beliefs about the USSR despite behavior inconsistent with their reputational beliefs (Holsti, 1967). In more recent memory, Serbia's obvious and horrific war crimes in Kosovo—clearly known to Russian elites at the time—were dismissed rather than integrated into any reputational beliefs by Russia, filtered instead through the lens of what Mendeloff (2008, 33) calls the "myth of Slavic Brotherhood." Finally, there is suggestive evidence linking helping behavior towards Ukraine (following Russia's invasion) to ingroup categorization among Europeans (Politi et al., 2023; see also Sinclair, Granberg and Nilsson, 2023).

These empirical anomalies dovetail with recent reviews: in a statement of what is missing from reputation research, Jervis, Yarhi-Milo and Casler (2021, 185-86) write that recent works neglect how beliefs about reputations might be "heterogenous. . . within and across states" and "stop short of observing or explaining. . . what factors are driving such dynamics." The exhortation of these conceptual reviews—highlighting unexplained heterogeneity in reputations—alongside the examples listed above suggest several puzzling facts: reputations do not always perfectly reflect behavior, the direction of the error seems to be systematic and there might be bias among observers in how behavior is integrated into larger sets of reputational beliefs.

Our argument is that one important source of variation in these inferences comes from *identity*, and in particular, the categorization of target states/actors into either the ingroup or outgroup.[1] Identity is "crucially important for understanding contemporary life" (Abdelal et al., 2006, 695) and "among the most normatively significant and behaviorally consequential aspects of politics"

---

[1]Of course, there are other sources of variation based on, e.g., the events and actions that generate their reputations in the first place (Clare and Danilovic, 2010; Weisiger and Yarhi-Milo, 2015; Yarhi-Milo, 2018).

(Smith, 2004, 304). If "...reputations and credibility are in the eye of the beholder..." (Jervis, 1989, 514), then our argument is that the lens of identity provides a powerful filter through which actors view events and update their beliefs about others.

We make both theoretical and empirical contributions in this paper. First, we provide a theoretical framework—based on social identity theory—to explain how identity dynamics affect both how we perceive other actors' "type" and estimate their likely behavior; in short, their reputations. Our argument focuses on group categorization—into ingroup/outgroup—and highlights two ways in which identity dynamics—and in particular, ingroup bias—affect reputations. First, we argue that categorization affects baseline judgments of other actors, leading to ingroup members getting an artificial boost from observers in their judgments of type and likely behavior. This improves upon current theories of reputation by providing a theoretical basis for judgments about reputation in the absence of any behavior (or in the presence of unclear patterns of behavior). Second, we argue that ingroup categorization has additional payoffs for target states, as it changes not only how actors *perceive* them, but the willingness of other actors to engage with them, help them, and defend them. Finally, we argue that the lens of ingroup identity affects how behavior by other countries—or information about them—is integrated into broader reputations. Ingroup states thus not only start off with better reputations but are more easily able to build reputations and pay fewer costs for reputation-inconsistent behavior compared to outgroups.

Empirically, we make two contributions. First, we field a descriptive survey on identity in international politics among the general public and two different samples of elites. Our survey results both provide useful information on which to build our experimental interventions as well as help to answer substantive questions about the boundaries of ingroups in international politics. From them, we learn that elites and the public converge in their beliefs about ingroups in world politics (Pearson's $r = 0.9$), that both groups hold similar reputational beliefs about other states and that there is a strong relationship between the belief a state is in one's ingroup and its ranking on valued attributes.

Our experiment builds on our descriptive survey and uses a repeated measures design to test implications of our theory in a setting where we "switch" a real country from ingroup to outgroup. Notably, we expand the scope of reputations from the traditional focus on reputations only for resolve, and elicit beliefs about several types of reputation, including financial, humanitarian and signaling. We find broad support for the notion that ingroup categorization affects states' reputations across many substantive domains, and also has significant benefits for observers' willingness to engage with the target state, increasing everything from the willingness to interact with and host its citizens, to respondents approving of trading with and even defending the country in wartime. Our design also allows us to pinpoint from which direction the causal effect of group categorization originates: we find that ingroup membership confers significant benefits relative to not being so categorized, but that categorization as an outgroup does not have any detectable effect relative to a neutral condition in our design. Finally, we find limited evidence that ingroup membership moderates the effects of past actions on reputation: reputation-building behavior—such as standing

firm in a crisis—from ingroups is significantly more effective than the same behavior from outgroup members.

1   Where Reputations Come From

Over the past few decades, a general consensus has formed around the core meaning of reputation—beliefs about a trait or behavioral tendency of an actor—as well as what questions are critical to understanding how reputations work. There are differences in focus, or emphasis, but *almost all work assumes that observable behavior is what generates and updates reputations*. That reputations are "informed" or "determined" by past behavior is in many cases simply part of the definition of what a reputation is (e.g., Dafoe, Renshon and Huth, 2014; Dafoe and Caughey, 2016; Renshon, Dafoe and Huth, 2018). In one canonical interpretation, reputations come simply from "observing behavior in context" (Tomz, 2007; Weisiger and Yarhi-Milo, 2015, similarly see reputation as being about product of past behavior). For others, there are theories specifying which *types* of behavior are most informative; for Crescenzi (2018), the focus is on how states learn from extra-dyadic interactions, while Lupton (2020) argues that early interactions for new leaders are particularly significant. Renshon, Dafoe and Huth (2018) show how actors update reputational beliefs—in particular, how they are differentially attributed to states and leaders—based on beliefs about which actor was responsible for a given observable behavior.

The fault lines in reputation research are thus rarely about what reputations *are*, but instead whether or not (and how) they "matter." In fact, most research on reputations is on what Kertzer, Renshon and Yarhi-Milo (2024) refer to as the "diagnostic" stage (whether judgments about reputation influence estimates of future/present behavior). In this debate, the key points of tension regard whether reputations factor in at all to current estimates of behavior, or whether they are overwhelmed by the current situation's stakes (and other features of the context).

1.1   Cracks in the facade

However, there are theoretical and empirical reasons to doubt at least the strong version of the "reputations as behavior" story in which reputations are formed and updated based almost entirely on past behavior. Empirically, we already know that observable behaviors alone are not sufficient to explain reputational inferences, and conceptually, our theories of reputation do not fully account for (systematic) variation across observers due to the psychological needs that past work suggests are critical in motivating behavior and perception.

Specific examples can serve to emphasize the first point. There is sometimes a large gap between what a reasonable observer *ought* to perceive in terms of a state's reputation if it was based only on behavior, and what we observe instead. Mercer (2023) highlights the example of European observers on the eve of the Russo-Japanese War in 1905. Observers in Europe were motivated towards accuracy by their security and alliance concerns, and had plenty of fodder—in the form of recent conflicts—on which to base their reputational judgments. And yet, they expected Russia

to easily defeat Japan, despite Japan's recent victory over China, its decades modernizing its military, qualitative advantages in military personnel and equipment, etc. Overall, Mercer (2023) documents the systematic racism that "distorted" assessments of Japan's "capabilities, interest and resolve." That the dramatic, public events that we tend to think affect reputations might be interpreted differently could strike one as surprising, yet even things that we might take for granted—whether a country won a war or suffered defeat, for example—are subject to the vagaries of post-hoc interpretations and motivated reasoning (Johnson and Tierney, 2003; Johnson, Tierney et al., 2006).

And while historical evidence suggests that observed behavior is translated imperfectly into reputations, other recent work suggests that the issue might be more systematic. Brutger and Kertzer (2018, 707) argue for a "dispositional theory of reputation in which individuals draw on their pre-existing foreign policy orientations to make reputational judgments." Even observing the same events, different "types" make different reputational inferences, because they care about different things: doves see threats as damaging to reputations (whether or not they succeed), for example, while hawks see the same threats as being beneficial to boosting reputations for toughness.

The second crack in the facade of the "reputations as behavior" story is conceptual. In fact, the gap between what reputations *would be* if they were derived solely from behavior and we observe in practice is one piece of a larger problem in the reputation literature in which we have paid short shrift to the actors making the reputational inferences. In one recent review, the authors note that the reputational beliefs that leaders hold are influenced, "like other beliefs by very human personal, psychological and political needs" that deserve far more attention than they have attracted thus far (Jervis, Yarhi-Milo and Casler, 2021, 189). If this recent call is to focus on psychological needs in our theories of reputation, then surely among the most powerful psychological needs one can imagine is the need for a positive and distinctive *identity*.

2　Identity and Reputations

We ask: what are the ways that identity (particularly through the categorization of others as ingroup or outgroup members) affects perceptions of other actors in world politics? Our answer is that identity dynamics—as reflected in social identity theory—affect both how we perceive other actors' "type" and our estimates of their likely behavior. Our particular focus here is on the aspect of SIT concerned with categorization into ingroups and outgroups and its attendant impact on how we perceive other actors in world politics.

While there are a "dizzying array" (Brewer, 2001, 115) of definitions of social identity theory,[2] a through-line in all the "inventions and re-inventions" of the theory has been the idea that "an individual's self-concept is derived, to some extent. . . from the social relationships and social groups he or she participates in" (Brewer, 2001, 117). Group membership provides people with a sense of where they stand with respect to others, a practical guide for how to act towards others and,

---

[2]As well as a host of related theories that overlap considerably. For example, on the subtle distinction between social identity theory and self-categorization theory, see Huddy (2001, 132).

finally, a lens through which to understand the behavior of others. Categorization into "us" and "them" is a "cognitive necessity" since one cannot begin to figure out how to act or what to think until they've "simplified the whir and buzz of our social environment" (Mercer, 1995, 241).

Of course, decades after the "sociological turn" in IR (Kowert and Legro, 1996; see also Katzenstein, 1996; Checkel, 1999), it is by no means novel to suggest the importance of identity. Older brush-clearing reviews (Tetlock and Goldgeier, 2000) implored scholars to take up issues of identity in IR and recent works (Kertzer and Tingley, 2018) suggest that they have begun to do so. Social Identity Theory has been particularly influential, for example, in the study of status, where the theory is used not to explain inferences about other actors but instead to provide a theoretical framework for understanding status-seeking behavior based on the desire for positive self-image (e.g., Larson and Shevchenko, 2003; Wohlforth, 2009; Murray, 2018). However, SIT has also seen use in theories of nation-building (Sambanis, Skaperdas and Wohlforth, 2015), EU expansion (Curley, 2009), threat perception (Gries, 2005; Rousseau and Garcia-Retamero, 2007) as well as broader theories of anarchy and self-help (Mercer, 1995) and rationality (Mercer, 2005) in IR. Examples abound from other subfields as well, most notably in American and Comparative politics, where a number of works have made use of social identity theory to understand partisanship (Greene, 1999; Bankert, Huddy and Rosema, 2017; West and Iyengar, 2020)

Two reasons stand out as explanations for why it's been commonly invoked in the past and why it might be suitable for theory-building in the domain of reputations. First is its distinction between interpersonal situations—governed by "personological," or individual-level variables—and group situations, in which behavior is determined by "category-based processes" (Brown, 2000, 746); the latter obviously analogous to politics, in which groups and group-based processes are critically important (Kertzer et al., 2022). Second are the conditions under which social identity is likely to exert the strongest impact on perceptions and behavior: when outgroups are seen as "competing for resources" and "when groups view the outgroup as having a history of tense relations" (Islam, 2014, 1782), two conditions that are ever-present in international affairs.

2.1 Ingroup Identity in International Politics

What Group Identities Matter in IR?

The first puzzle to address is determining the boundaries of identity in world politics. In the vast majority of work on social identity theory, identities are either created on the spot based on arbitrary distinctions—the famous "minimal group paradigm" (Otten, 2016)—or are based on categories such as gender and race, backed by decades of research on identity salience in daily life (Cikara and Van Bavel, 2014, 247) . We also have by now a fairly good sense of contextual factors—e.g., whether an identity is under threat or elicited in the presence of outgroups (Cuhadar and Dayton, 2011, 275; McGuire et al., 1978)—that make given identities more salient.

Yet, we are on far less sure ground when it comes to understanding identity categories in politics. Early research on multiple identities suggested that we might expect predictions from social identity theory to be *most* applicable to "ethnic, religious, political identities because they

are more 'collective in nature' than other individual aspects of identity" (Deaux et al., 1995, 286).[3] And more recently, we've amassed compelling evidence that partisanship functions as a social identity in *domestic* politics (Greene, 1999; Huddy, Mason and Aarøe, 2015; West and Iyengar, 2020), but what categories would be salient in international politics is far less clear.

In conceptualizing international identity, we take our cue from Abdelal et al. (2006) in focusing on "relational comparisons" with other groups; in essence, "defining an identity group by what it is not, i.e., the way it views other identity groups, especially where those views about the other are a defining part of the identity" (Abdelal et al., 2006, 696).[4] This is particularly important given the history of adoption of SIT in international affairs, which has tended to assume that *any* outgroup represents the "other"—that there is "a black-white dichotomy between 'us' and 'everybody else'" (Hymans, 2002, 10). In fact, only relevant comparison outgroups are likely to trigger the biased intergroup evaluations described in more detail below, begging the question: what are the relevant outgroups in international politics? Our exploratory research question— listed alongside our hypotheses in Figure 1—is thus, **what group identities are salient in international politics?**

2.2 How Categorization Shapes Reputations

While the boundaries of group identity in world politics are yet unknown, the implications of that group identity are more clear. Primary among them is ingroup bias (Brown, 2000, 747). Individuals have a deep-seated need to see themselves—and the groups they belong to—in a positive light, leading them to evaluate their own group more favorably than other groups (Cuhadar and Dayton, 2011, 274). Summarized by Druckman (1994, 48), ingroups are "perceived as better, friendlier, more competent, and stronger than other groups." And, critically, the "act of categorization" has a significant effect on not just perception of others, but our behavior (Cikara and Van Bavel, 2014, 247).

A basic understanding of how ingroup bias operates generates three empirical implications. First, **categorization as an ingroup member will positively influence observers' beliefs about that target's "type," attributes and likely behavior; i.e., their reputation (H1)**. Observers—whether they are elite officials or members of the general public—will see ingroup members as more moral and honest, more capable, more likely to act in a resolved manner, and to honor their commitments. This first empirical implication is a direct translation of ingroup bias to international affairs, and indeed, it would be surprising given the empirical foundation for ingroup bias in psychology if this was not the case. The challenge for this implication derives instead from constructing a fair test for it, given that we are on shaky ground in understanding the "sources of

---

[3]See also Huddy (2001, 148), who notes that "It is hard to believe that longstanding political identities linked to major ideologies or political parties would exhibit the same high level of fluidity as an artificial identity created in the lab."

[4]This is one of three "forms" outlined by Abdelal et al. (2006) in their seminal work on measuring identity: constitutive norms (formal and informal rules that define group membership), social purposes (goals that are shared by members), cognitive models (worldviews) and relational comparisons with other social categories. Our interest here in the latter type.

Figure 1: Research Questions and Hypotheses

social identity" and ingroup boundaries in international politics (Huddy, 2001, 128).

The second implication of our ingroup bias argument is that identity impacts how behavior—or more generally, information about other states—is understood and incorporated into beliefs about targets. In what we see as the traditional model (the "reputation as behavior" story), observers witness behavior by other states or otherwise learn information about them and use that to update their beliefs about that state; whether about their attributes, likely future behavior, whether they are "like us" or not, or willingness to cooperate in the future (e.g., Dafoe, Renshon and Huth, 2014; Powers, 2023). Our H2 tests one implication of this argument. Specifically, we argue that **behavior witnessed by observers will be judged differently depending on whether the target of inference is considered to be in the observers' ingroup or not.**

Our final hypothesis addresses the behavioral consequences of ingroup bias in reputational judgments. While Here, a large body of research in cognate fields suggests that shared ingroup identity directly affects helping behavior (e.g., Levine et al., 2005) on an interpersonal level and in more diffuse ways like the willingness of individuals to consume goods from culturally distinct countries (e.g., Bankert, Powers and Sheagley, 2022). Some more recent work extends the theory directly to politically-relevant behavior as well (Distelhorst and Hou, 2014). We predict that **observers will be more willing to engage with states perceived to be ingroup members**. We focus here on a wide range of outcomes, from willingness to interact socially with citizens of the target state, to decisions related to financial aid and even willingness to fight alongside the target state.

2.3   Relationship to related theories of reputation

Given the number of existing theories of reputation, it's worth pointing out where our argument about identity diverges from other work (as well as where it generates overlapping predictions). One important contrast is how our argument relates to work in the "reputation as behavior" genre, in which they key prediction is that reputational beliefs will be shaped by past behavior. If these theories (e.g., Tomz, 2007; Weisiger and Yarhi-Milo, 2015) predict that, e.g., $i's$ resolved behavior will overall lead others to see i as more resolved, ours does as well, though with a twist.

Our argument does not ignore behavior but rather predicts that categorization will shape (read: moderate) the effect of behavior on reputations. Resolved behavior will still affect reputations in the "correct" direction, but the reputation of ingroup members will benefit from good behavior disproportionately compared to outgroup members. Where our theory provides the most contrast is in the *absence of any observed behavior*, where extant theories of "reputation as behavior" do not provide much traction, and where ours generates clear predictions about the factors—ingroup categorization—that will shape reputations.

Another important contrast is with Mercer's (1997) psychological theory of reputation, in which reputation updating is conditional on attributions of responsibility. In simple terms, people interpret behavior as due to dispositional factors when it involves an "outgroup's undesirable behavior" (and due to situational factors when it involves desirable behavior) (p. 9). Because only dispositional attributions can generate or update reputations, "adversaries can get reputations for having resolve...while allies can get reputations for lacking resolve," but not vice-versa (p. 10). Three principal differences emerge from our argument, one conceptual and two empirical: (1) Conceptually, our argument does not require the "input" of behavior in order for ingroup categorization to have an impact on reputation; (2) for Mercer, all other countries are "outgroups," whereas for our theory, other countries may be part of an actor's ingroup; (3) our argument predicts that, within ingroups, resolved behavior by a target state will positively affect peoples' estimates of their reputation (not so for attribution theories of resolve, in which desired behavior by allies is ignored). The section below details how we test our argument as well as the diverging predictions from related theories of reputations.

2.4 Research Design

The brief discussion above highlights two priorities of any empirical investigation of identity in international politics, one exploratory and one causal. The first is investigating the *boundaries* and content of international identity. There is at least suggestive evidence that doing so is possible, and that the kinds of identity dynamics that operate at the domestic level of politics or in daily life are active in international politics. For example, "longstanding political identities" ought to be stronger and less fluid than, at least, artificial identities created in a lab (Huddy, 2001, 148), and the more "collective nature" of political identities might contribute to its import (Deaux et al., 1995, 286).

Our primary interest here is in exploring which of the many possible identities in international politics might be most salient. Reckoning with the question of "identity choice" when so many are possible is one of the most pressing issues required in order to "adapt social identity theory to political phenomena" (Huddy, 2001, 131). Addressing these substantive questions helps us understand how international identity operates—answering the call by Reicher (2004) to incorporate some of the richness from the psychological theories we adapt—as well as lays the groundwork for our investigation of how categorizations impacts reputation.

The second task is to link identity—particularly perceptions of other actors as being in one's

Question	Theoretical Estimand		Empirical Estimand	Method
	unit-specific quantity	*target population*		
What identities matter in international politics?	descriptive: who is in the U.S.' ingroup and what are their reputations?	U.S. population of (a) citizens (2) elites	mean in samples of (1) N=1,500 LucidTheorem respondents, using census quotas for region, gender, race and age (2) N=650 elites recruited by TRIP	Survey
What is the effect of ingroup categorization on reputations?	causal: difference in reputation (Y) for i when it is categorized as ingroup ($d = 1$) vs. outgroup ($d = 0$)	adults (no geographic restrictions)	average treatment effect of ingroup encouragement for Turkey (compared to outgroup and neutral) in U.S. online convenience sample	Survey Experiment

Table 1: Theoretical and Empirical Estimands

ingroup or outgroup—with reputation-related beliefs and behaviors. We expect to find evidence that individuals see ingroup members in a more positive light than outgroup members, with higher ratings for valued attributes and lower ratings for negative ones. We also expect that because biased ingroup evaluations will affect beliefs about the attributes and "underlying type" of other actors, it will carry implications for judgments about likely behavior. An additional empirical implication of our argument is that ingroup bias will act as a filter through which behavior is observed, moderating the impact of behavior on reputations. Finally, consistent with a wide literature on ingroup bias and behavior, we expect to see that ingroup categorization will increase the likelihood of cooperative behavior toward the target state.

Table 1 captures our estimands, using a simple version of the framework suggested by Lundberg, Johnson and Stewart (2021). Theoretical estimands are the "questions outside of the data" and are a combination of a (1) unit specific quantity (a realized or potential outcome) and (2) a target population. The unit-specific quantity clarifies whether the object is to make descriptive or causal inferences, and the target population addresses the question: over whom or what do we aggregate that unit-specific quantity (Lundberg, Johnson and Stewart, 2021, 534)? The empirical estimand takes into account real world constraints and focuses only on observed quantities. The value in explicitly stating these quantities is greater clarification about what research design is optimal, what sources of data ought to be used, and most importantly, what assumptions we must make in order to connect our theoretical to our empirical estimands. The second row of Table 1, for example, shows that while the target population for our theory has no geographic restrictions, our data will come from an online sample in the U.S., and while our theory is about broader effects of ingroup categorization, our particular design will estimate those within the context of a series of treatments about one particular country.

Explicitly detailing our theoretical estimands also helps connect the target of our investigation to an outcome through a particular method and set of identifying assumptions. Given that our first question implicates a descriptive quantity—perceptions and beliefs about the U.S. in relationship

to other countries—we turn to survey methods fielded in a U.S. sample of elites and citizens. And given the limited evidence for sensitivity bias in most contexts (Blair, Coppock and Moor, 2020), we use a combination of direct and indirect questions. Our second set of questions about the effect of ingroup categorization on reputational beliefs is causal, and implicates the difference between what the reputation for a state, i would be under two different possible states, when it is categorized as either an ingroup or an outgroup. To investigate this set of questions, we turn to survey experiments.

3   The Content of Group Identity in World Politics

We begin with our descriptive surveys fielded on the public and foreign policy elites. Our goal is to identify—for these two important populations—who is in the ingroup, who is in the outgroup, and what marks them as such. These surveys thus directly address our exploratory research question: what group identities are salient in international politics?

In addition to answering a substantively important question in its own right, this exercises also yields two important practical benefits. First, we are able to use insights from these initial descriptive surveys to motivate the design our experimental studies that follow. Indeed, determining the content and boundaries of identities relevant to international politics is a necessary precursor to any effort to manipulate those identities in an experiment. Second, by fielding parallel surveys on elites and the public, we can also speak to the degree of convergence between attitudes of the public and foreign policy elites on questions of identity and reputation. Convergence in this particular domain would better allow for the use of lower-cost general population samples as stand-ins for foreign policy elites as we assess the causal role of group identity on reputational inferences in world affairs (Kertzer and Renshon, 2022).[5]

We fielded our descriptive surveys in Fall 2022.[6] Our U.S. general population sample ($N \approx 1,500$) was recruited by LucidTheorem, who used census quotas on region, gender, race and age to ensure that our sample included a broad cross-section of the public.[7] Our foreign policy elite sample ($N \approx 650$) was recruited by the Teaching, Research, and International Policy (TRIP) Project (Avey et al., 2023). Our sample of US elites consists of a combination of academic faculty who work

---

[5]Keeping in mind, of course, that the two empirical components of our paper are estimating two different quantities—attitudes and responses to treatment—and so convergence on one does not necessarily imply convergence on the other (Kertzer, 2022).

[6]Appendix §A.2 contains the full instrument fielded on the Lucid sample in early September, 2022. A slightly shorter version, also in the Appendix, was fielded on the TRIP sample.

[7]In the appendix, we compare the distribution of important demographic variables in our sample to a recent national benchmark survey, the 2020 Congressional Election Study, and find that, like other studies employing online convenience samples (Huff and Tingley, 2015; Berinsky, Huber and Lenz, 2012), our respondents are somewhat younger, lower income, and more liberal than the general public. Coppock and McClellan (2019) find that "demographic and experimental findings on Lucid track well with national benchmarks." Ternovski and Orr (2022) document a decline in attentiveness of Lucid/Cint respondents after 2020. We deal with this by using strategies similar to those outlined in (Berinsky et al., 2021). First, we use a relatively easy to pass attention check to screen out the most inattentive respondents (so-called "speeders") at the very beginning of the survey. Second, we include a number of attention checks (in a variety of formats) throughout the survey, allowing us to generate a continuous measure of attentiveness so that we may study how responses vary as a function of attention level.

on International Relations and U.S. foreign policymakers.[8] The sample includes individuals who work(ed) in offices in the White House, Defense Department, State Department, Treasury Department, US Agency for International Development, among others over the period 2000-2020 and all faculty currently appointed at U.S. colleges and universities in a political science or international relations department who study issues that cross borders. In all, 215 policy elites and 429 faculty members answered our survey. We present results from the policy elite sample below, but note that all results here hold among our faculty member sample as well.

Survey Questions

In our surveys, we presented respondents a list of 10 countries[9] and asked them to categorize the countries as either 'us' or 'them.'[10] This exercise allows us allows us to identify precisely *who* is in the ingroup and outgroup for each population. Moreover, the shared attributes of the members of the resulting groups will provide important clues as to how individuals distinguish ingroup from outgroup members. Following our group categorization task, we selected three countries at random from the ten the respondent had initially categorized. We asked respondents to indicate their level of agreement with six statements about each of these three randomly selected countries.[11] Among several other dimensions, we asked about agreement with statements that the country in question is trustworthy, likely to stand firm against enemies, likely to honor commitments, likely to pay its debts, and similar to the United States.[12]

3.1 Three Lessons on International Identity

We take three principal lessons from our descriptive surveys. First, the public and policy elites tend to agree on who is in the country's ingroup and these countries tend to have shared attributes (e.g., geography, culture, religion, language, and/or regime type). Second, elites and the public hold similar reputational beliefs about other countries. And third, both among the elites and the public, "ingroup-ness" is correlated with positive traits. We discuss each of these in turn.[13]

---

[8]The sample frame construction follows closely the procedures outlined for previous iterations of this survey in Avey and Desch (2014) and Avey et al. (2021).

[9]Randomly drawn a group of 26 countries: the 25 most powerful countries as measured by the Correlates of War "composite index of national capabilities" plus Israel.

[10]Our approach here is adapted from Pinto et al. (2020) and White and Dahl (2007) who study the role of identity in consumer purchasing decisions.

[11]The question read, "We're interested in perceptions of other members of the global community. Below, we ask for your level of agreement with a number of statements about countries and groups of countries around the world. Please give your impressions of the country or group, whether or not you know much about them. The country of X is..."

[12]The response scale for these questions was continuous and ranged from 0 (no agreement at all) to 100 (total agreement). We selected these dimensions because because past work has shown that they are central to our understanding of conflict and cooperation in world affairs: Kydd (2007) documents the important role that trust plays in cooperation between would-be adversaries, Tomz (2007) shows that reputations for debt service facilitate international financial relations, Kertzer (2016) links reputations for resolve to success in crisis bargaining, and Chen, Pevehouse and Powers (2023) find that reputations for honoring commitments are key to understanding the democratic advantage in the
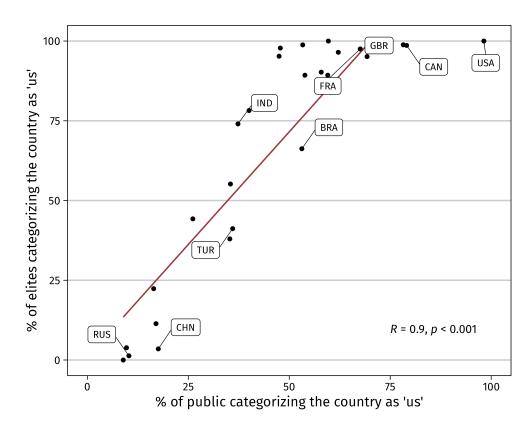
Figure 2: **Elites and the public agree on who is in the ingroup.** We asked respondents to categorize countries as either "us" or "them." The scatter plot shows that countries that elites were most likely to categorize as "us" were also categorized as "us" by the public.
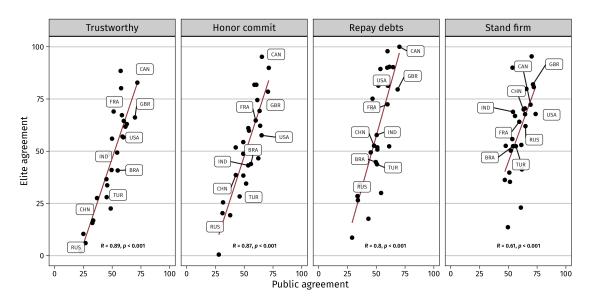
Figure 3: **Elites and the public hold similar reputational beliefs.** We asked respondents to indicate their level of agreement with a variety of statements about different countries.

Elites and the public perceive international identity similarly

The first substantive lesson from our descriptive surveys is that foreign policy elites and the public agree on who is in their ingroup. We calculate the proportion of the public and foreign policy elites who categorized each country as in the "us" group and plot these quantities against each other—elites on the vertical axis and the public on the horizontal axis—in Figure 2.

Two features of this plot are worth remarking on. First, those countries that are more likely to be categorized as "us" are distinguished from those that are more likely to be categorized as "them" in that they are similar to the United States one or more social or political dimension (e.g., geography, culture, ethnicity, religion, regime type, and/or alliance status). Indeed, the six countries with the highest "us" placement (Canada, UK, Italy, France, Germany, and Spain) are all relatively wealthy democracies with predominantly white and Christian populations and are NATO members. We use this result—that shared political, ethnic, religious, and/or cultural features are important markers of group identity—to motivate the design of our identity manipulation below. Second, there is a very high degree of convergence between the two samples, with a Pearson correlation coefficient of .9. Taken together our initial results suggest that social, cultural, and political *similarity* may be an important marker of group identity in world affairs and that is so for both the public and foreign policy elites.

Elites and the public hold similar reputational beliefs about other states

The second lesson that emanates from our data is that our respondents hold similar reputational beliefs about countries around the world. We take the mean level of agreement that the country in question is trustworthy, likely to honor commitments, likely to repay debts, and likely to stand firm in a crisis (for each sample) and plot the resulting quantities from each sample against each other in Figure 3. Here, again, we see remarkable convergence between public views and those of elites. Elites and the public not only agree on who is in the ingroup and what defines them as such; they also hold similar reputational beliefs about those actors. Agreement ranges from an r of 0.61 for "stand firm" to between 0.8-0.9 for "'trustworthy," "honor commitments," and "repay debts."

Elites and public associate positive traits with "us-ness"

Finally, there is a strong relationship between ingroup-ness on the one hand and valued attributes and traits on the other among both the public and elites. For each sample and each reputational belief in Figure 4, we calculate the Pearson correlation coefficient between the % of the sample categorizing a country as being in the ingroup and ratings on that particular trait/behavioral tendency. We have ordered the reputational belief panels such that the highest correlation between ingroupness and reputational belief is on the left and the lowest correlation is on the right. Doing so reveals yet another dimension of correspondence between elite and public views: we showed above that elites and public converge not just in their views on group membership and the reputational beliefs they hold about those actors. Here, we show additionally that the relationship between ingroupness on the one hand and positive reputational beliefs on the other varies across dimensions in similar ways among the two samples. For both elites and the public, trustworthiness appears linked strongly directly to ingroupness but this link is not quite as pronounced when it comes to perceptions of resolve (and less so for elites compared to the public).

3.2 Informing our causal research design

Our descriptive work above provides a valuable foundation on which to build our experimental designs along several dimensions. First, the correlation between ingroupness and valued attributes provides empirical motivation to investigate the possible causal relationship more directly using experimental methods. We come away from our descriptive surveys with significant correlational evidence from two samples showing that ingroup members get a reputational subsidy, directly in line with our theoretical predictions above. Notably, however, ingroup-ness did not always correlate in the same strength with all reputational inferences. Perceptions of resolve, for example, seemed less related to ingroupness than the other dimensions of reputation and so any effort to study

_____

political economy of cooperation.

[13]Below we show analyses based on data from our public and policy elite samples, omitting data from IR scholars. Appendix §A.3 shows similar graphs that include responses from our IR scholar sample. The results are nearly identical.
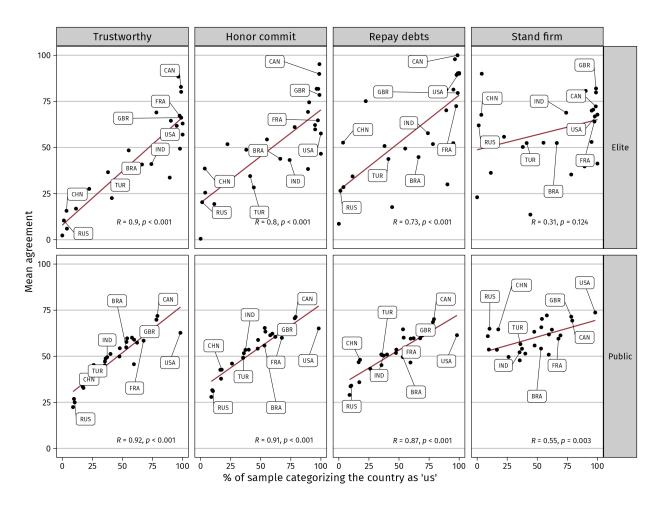
Figure 4: **Positive traits are correlated with "us-ness."**.

the identity and reputation experimentally must take into account the multidimensional nature of reputations in international relations.

Second, our survey results illuminated not just who the U.S. public and elites see as in their "international ingroup," but which states find themselves in a kind of liminal space between the ingroup and the outgroup. Turkey, for example, was consistently in the middle third of the "ingroup" distribution. Turkey and other states like it may, from the perspective of our respondents, not fit cleanly in the ingroup or the outgroup. In that sense, their identities may be more malleable than more consistently categorized states (e.g., the United Kingdom or Iran). As such, Turkey may be a good candidate for use in experiments in which we try to "nudge" respondents' perceptions of a state's group membership in one direction or another direction.

Third, while much past work on social identity has emphasized the important role of shared attributes or similarity as an "input and an outcome of the social categorization process" (Spears, 2021, p. 372), our descriptive survey results help us identify *which* attributes are likely to be salient to members of the public and elites in this context. We use these results to inform the design of our experimental manipulation. In particular, we draw on the fact that those states most likely to be categorized as members of the ingroup (outgroup) tended to be similar (dissimilar) to the United States on a number of salient political and cultural dimensions, including geography, regime type, alliance status, language, and/or religion. As we detail below, we attempt manipulate respondent perceptions of Turkey's group membership by shifting perceptions of similarity on these dimensions.

Finally, the elite-public convergence in our initial surveys suggests that we are on firmer ground in fielding our experiments on public samples not just as an initial empirical foray or because we think the public's beliefs matter in this area (though we do), but because they are a plausible stand-in for elites whose perceptions and judgments are of central importance. We use all of these descriptive insights to inform the design of our experimental manipulations below.

4   Causal Research Design

Our goal in the causal research design is to estimate how being perceived as being in a common ingroup (relative to an outgroup) affects observers' perceptions of another country—beliefs relating to reputation—as well as their willingness to engage (at different levels of intensity) with that target state. Our second goal is to learn about how the effect of shared ingroup-ness moderates the impact of observed behavior on reputations. Put simply, does "bad" behavior affect ingroup members differentially compared to outgroup members?

In our pilot, we manipulate respondents' perceptions that a target country (Turkey) is in their shared ingroup or their outgroup, and estimate the effect of that intervention directly on reputational judgments. We also include a behavior manipulation to allow us to investigate how shared ingroup-ness moderates the effects of behavior, helping to answer substantive questions about whether the bad behavior of ingroup members is discounted relative to outgroups. In this study, we also provide additional information about Turkey to fix potential confounders and maintain

information equivalence as much as possible.[14]

The target country is Turkey, which was chosen based on our survey data, where we established that U.S. respondents placed Turkey squarely in the middle third of the distribution for countries that are in the U.S.' "ingroup." The implication of that is that it is a context that is amenable to nudging in one direction or another. It's also a country about which Americans know comparatively little, so there are fewer concerns about strong priors or pre-treatment.

4.1 Pilot: A Priming Design

Our main pilot—depicted in Figure 5 and described in detail in Appendix B—approaches identity and shared in-group-ness directly by manipulating respondents' perceptions of Turkey as an ingroup or outgroup member. The design utilizes a repeated measures structure (Clifford, Sheagley and Piston, 2021) to improve precision in which all subjects complete DV measurement at three points: upon entering the study (T_1), after the identity manipulation (T_2), and then, finally, following the behavior manipulation (T_3).

Our Identity Treatment
Our identity treatment is depicted in Figure 6. Fundamentally, the treatment operates through a combination of (1) information provision, (2) a graphical depictions of Turkey's leader and it's geographic location relative to its neighbors, and (3) respondent participation in a brief thought-listing exercise. No deception is used. Instead, different aspects of Turkey are emphasized in the ingroup/outgroup arms. Respondents in the ingroup, for example, read about Turkey being included in maps of Europe in many atlases, its application to join the EU and other features that emphasize its commonality with the United States, including language, popular culture and religion. Respondents see a graphic showing how close Turkey is to EU countries and emphasizing its place in Western Europe as well as a photo of President Erdogan in a western-style suit.[15] Finally, they are asked to think about the things that the U.S. has in common with Turkey and list several of them.

---

[14]Though, it's more complicated, because, actually, many possible confounders are also signals of identity and thus may function as confounders *or* causal pathways through which identity affects the DV.

[15]In the OUTGROUP arm, respondents saw a photo of President Erdogan in more traditional Arab clothing and a kefiyeh. This manipulation was inspired by the work of Devra Pager (2003) and more recently, a very similar treatment used by Fernández-Reino, Di Stasio and Veit (2023). The photo of Erdogan was taken during a 2012 visit to a refugee camp for Syrians in Turkey.
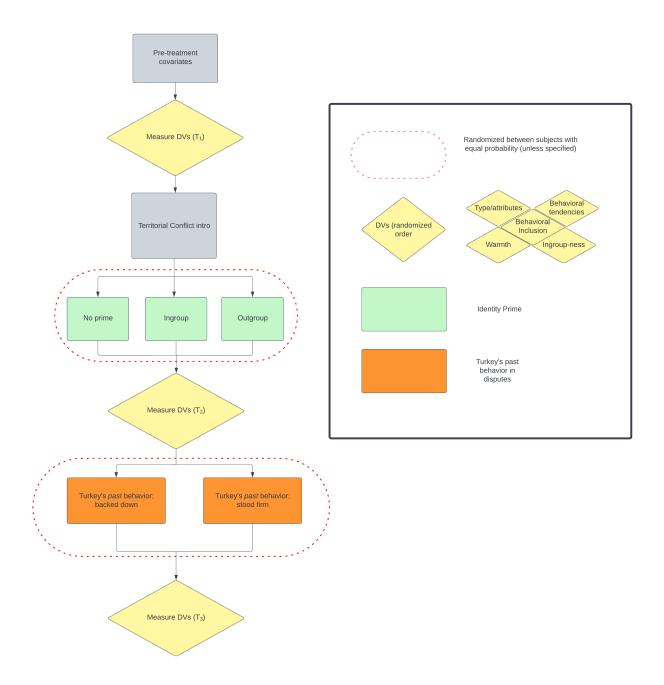
Figure 5: Consort Diagram for Pilot.

Ingroup: Some background about Turkey: The country of Turkey is included in maps of Europe in many atlases. It is officially known as the Republic of Turkey. The people of Turkey are quite familiar with American popular culture. Disney+, Netflix, and YouTube are quite popular in Turkey and Instagram is among the most popular social media services there. It has applied for membership in the European Union an organization designed to promote cooperation and cultural exchange between European states. Turkey shares a land border with two EU countries, Greece and Bulgaria. It takes less time to travel by air from Turkey's capital city to London, UK than from Washington, DC to London, UK. The most recently elected president of Turkey is President Erdogan (pictured below). The second most common religion in Turkey is Christianity and in many tourists spots English is spoken.

Outgroup: Some background about Turkey: The country of Turkey is included in maps of the Middle East in many atlases. It recently changed its official name to Türkiye. The people of Turkey consume a lot of popular culture from their region. One of the most popular streaming services in Turkey is BluTV, founded in Istanbul, which features lots of media from Turkey and around the region. It is a member of the Organization of Turkic States and TÜRKSOY, two organizations designed to promote cooperation and cultural exchange between Turkic peoples across Central and Western Asia. Its shares a land border with Iraq, Iran, and Syria. It takes less time to travel by air from Turkey's capital city to Iran's capital than from Washington, DC to London, UK. For the past 10 years, the leader of Türkiye has been Recep Tayyip Erdoğan (pictured below). The most common religion in Turkey is Islam and the most common language is Turkish.

Think for a moment about the parts of Turkey that are most similar to the United States. Please list four of them in the open response boxes below.

Think for a moment about the parts of Turkey that are most different from the United States. Please list four of them in the open response boxes below.

Figure 6: Identity Treatments

Our Outcomes

At the start of the survey (T_1), following our group identity manipulation (T_2), and then finally after our past behavior manipulation (T_3), we measure several quantities, listed below:

1. Warmth

2. Ingroup-ness

   Ingroup Categorization (Manipulation Check + placebo)

3. Traits/Attributes

4. Behavioral Tendencies

   Reputation (What respondents *think* about Turkey)

5. Support for Turkey battery

   Willingness to Engage (how respondents would *act* towards Turkey)

Our manipulation checks combine measures of warmth with a question about the extent to which Turkey—alongside a number of placebo countries (UK, France, Iran)—is in the respondent's ingroup or outgroup using a question adapted from our 2022 descriptive surveys.[16] Our primary substantive interest is in *reputations*, which are commonly agreed upon as comprising both beliefs about the traits/attributes of other actors, as well as their behavioral tendencies (Dafoe, Renshon and Huth, 2014, 372: see also Jervis, Yarhi-Milo and Casler, 2021). We ask about attributes or traits (tough, morally pure, trustworthy) as well as behavioral patterns (likely to stand firm, honor commitments, win wars, do the things it says it will do, repay its debts and accept refugees). Notably, we expand the scope of reputations from the traditional focus in the literature on reputations for resolve, eliciting beliefs about several other types of reputations including financial reputations (repayment of debts), humanitarian reputations (willingness to accept refugees), and signaling reputation (trustworthy and likely to do what they say). While we anticipate that our argument would apply to other dimensions of reputation as well, we chose these dimensions because of their central importance to theories of conflict and cooperation in international relations (e.g., Tomz, 2007; Kertzer, 2016; Chen, Pevehouse and Powers, 2023; Kydd, 2007).

While our reputation questions focus on how respondents *perceive* the target country—its attributes, characteristics and estimated future behavior—another set of questions focuses on the consequences of theses perceptions. More specifically, we ask a series of questions designed to investigate how respondents would *act* towards the target, broadly under the category of "willingness to engage."[17] These behavioral inclusion questions are adapted from similar questions commonly used in the context of intergroup relations and attitudes towards outgroups, where they focus on willingness to have someone from a racial or ethnic outgroup as a boss, neighbor, or spouse (Sniderman, Hagendoorn and Prior, 2004; Mason, Wronski and Kane, 2021). We thus selected our

---

[16]"In discussing world affairs, people sometimes use words like "us" and "them": "us" to describe our country and other countries that are like us in some way, or part of our group, and "them" to describe ones that are not like us and not part of the same group. Where would you place the following countries along an "us" to "them" continuum?

[17]Willingness to: allow consumer goods from Turkey to be sold in the U.S., allow tourists from Turkey to enter U.S., allow Turkish citizens to live and work in the U.S., allow Turkey citizens to become U.S. citizens, send money to help Turkey economically, send military supplies to Turkey, send U.S. military personnel to help defend Turkey.

outcome measures to speak to the role of reputation not just in relations between states but also the "transnational" relations of individuals, organizations, and firms. The importance of the former, of course, results directly from states ability to greatly affect the security and prosperity of large numbers of people, while the importance of the later stems from the "sheer number of exchanges" between the myriad smaller actors that constitute, in literal terms, the bulk of the international relations that occur on a daily basis (Findley, Nielson and Sharman, 2013).

4.2   Pilot Logistics

We pre-piloted the treatment to evaluate its plausibility in October 2023 on LucidTheorem on a sample of 200 respondents.[18] This pre-pilot of the identity treatment established that the ingroup treatment moved our manipulation checks in the correct direction, increasing the warmth felt towards the target country (Turkey) and the probability that respondents categorized Turkey as "us" (compared to "them") but not doing either for the placebo countries (UK, France, and Iran). This helps establish the efficacy of our treatment in manipulating our central concept (ingroup categorization) as well as helping to preempt concerns that the ingroup treatment, for instance, is simply changing how respondents feel about all foreign countries, or doing something in particular for countries like Iran that are located in roughly the same geographic region.

We fielded our full pilot on LucidTheorem (N=2,039) in October 2023.[19] In the survey, we used two pre-treatment checks to screen for inattentive respondents. Detailed information about pass rates and procedures in the Table 5 in the Appendix. We screen for inattention because Lucid has acquired a reputation for inattentive respondents, though that is balanced out by a recent finding that Lucid provides "the closest to representative probability samples on demographic and attitudinal representativeness" (Stagnaro et al., 2024). All our attention checks are pre-treatment except for one one additional post-treatment attention check in order to ensure that the "remainers" paid attention throughout the study (but do not use this information to censor the sample; Lo, Renshon and Bassan-Nygate, 2023). Ninety-eight percent of respondents who passed our first two attention checks also passed our end-of-survey attention check.

5   Ingroup Identity, Reputations and Helping Behavior

Our first goal here is to establish whether or not our identity manipulation works as anticipated, affecting relative levels of warmth, social distance, and othering. Conditional on our manipulation working, we wish to answer three subsequent questions. First, do reputational perceptions vary as a function of group identity and, if so, is there a bias in favor of the ingroup (H1)? Second, does the effect of new information about an actor's behavior depend on whether the actor is viewed by

---

[18]Results are depicted in Figure 14 in Appendix C. Obviously, this is not large enough to generate sensible power estimates. The purpose here was simply to identify potential problems and evaluate the basic plausibility of the treatment in moving our manipulation checks in the hypothesized direction.

[19]All studies were approved by the UW-Madison IRB (#2023-1092) and the University of Georgia IRB (PROJECT00006273).

the respondent as a member of the ingroup or not (H2)? Finally, does that ingroup bias carry over to affect not just perceptions of target states, but willingness to engage with that state and its citizens (H3)?

Our dependent variables are all on a 0-100 scale, with higher values indicating more "desirable" levels of each measure (e.g., more resolved, more honest, more like us). We estimate the effect of each of our group identity treatments relative to our control group using OLS, regressing each dependent variable—at T_2 (to study main effect of group identity) or T_3 (to study the conditional effect of past behavior)—on a set of treatment indicators and a set of conceptually relevant pre-treatment measures.[20] We plot the resulting estimates (and 95% confidence intervals in the figures below. Full regression results tables are in the Appendix.

Social Categorization Causally Affected by Ingroup Treatment

We first assess whether our treatment worked as intended by considering both our quantitative manipulation checks and text data from the thought-listing component of the treatment. The goal of the thought listing exercise is to verify that respondents completed the task in good faith and returned patterns consistent with the contrasts that we highlighted in our manipulations. We processed the raw responses, removing stopwords and punctuation and stemming the remaining words. We tokenize the responses into single words and bigrams (unique combinations of two word stems) and calculate the relative frequency of each token in each manipulation condition. We present the ten most frequently used word and bigram stems in each condition in Table 2. The results are consist with respondents addressing the task in good faith and cuing on precisely those features which we thought would shift perceptions of community membership (e.g., religion, culture, geography, and language).

Our manipulation checks (Figure 7) provide further evidence that our treatment worked as intended. Overall, our treatment succeeded in shifting the social categorization of Turkey: respondents felt warmer towards Turkey by 10.3 points on our 0–100 scale (95% CI: 10.6, 12.2; $p < .000$) and were more likely to categorize Turkey as being in the "us" group in international politics by 3 points on our 0–100 scale (95% CI: 2.54, 5.27; $p = .011$). Our inclusion of a neutral CONTROL group (no identity prime) allows us insight into where the "action" is, suggesting in this case that the INGROUP treatment is significantly affecting respondent perceptions relative to the CONTROL, while the OUTGROUP condition is indistinguishable from the CONTROL arm.[21] Results also show

---

[20]For example, we control for our pre-treatment measures (T_1) of Turkey's reputation when estimating the effect of treatment on reputation. Similarly, when estimating the effect of treatment on warmth and "us-ness," we control for our baseline pre-treatment measures of the same quantities. Results are qualitatively similar without our pre-treatment controls, but less precisely estimated, which is of course the point of this approach. See Clifford, Sheagley and Piston (2021) on the efficiency benefits of this approach.

[21]This might be driven by floor effects: i.e., respondents might feel quite "cold" towards Turkey to begin with such that there is no room for the outgroup treatment to work. Given that the mean level of warms and "us-ness" in each control group was in the low 40s, this is perhaps not so plausible. Alternatively, it could be that it is, in general, easier to nudge social categorization towards ingroups rather than outgroups (or in easier in particular cases such as Turkey). Notably, we do not find evidence consistent with a "rebellion" against the treatment in which some respondents felt warmer towards Turkey as a result of the outgroup treatment, potentially offsetting those who

Word stems

Ingroup condition		Outgroup condition	
Word	Pct. of Respondents	Word	Pct. of Respondents
christian	46.2	religion	48.6
english	46.0	languag	34.6
cultur	35.4	islam	29.4
religion	24.6	cultur	20.4
presid	24.2	border	19.2
media	20.8	countri	18.7
speak	20.7	leader	17.0
social	17.6	turkish	14.0
american	16.7	iran	12.6
spoken	15.3	turkei	11.4

Bigram stems

Ingroup condition		Outgroup condition	
Bigram	Pct. of Respondents	Bigram	Pct. of Respondents
speak english	21.5	middl east	14.7
social media	21.3	common religion	13.4
pop cultur	12.3	main religion	8.5
american cultur	9.8	land border	8.1
popular cultur	8.6	common languag	7.4
elect presid	5.8	speak turkish	6.5
stream servic	5.8	stream servic	5.8
common religion	5.4	iraq iran	5.1
land border	5.0	territori disput	4.1
english spoken	4.8	share border	3.2

Table 2: Word and bigram stems from the thought listing excercise.

that our INGROUP and OUTGROUP treatments do *not* generally shift perceptions of other countries used as placebos. The one exception is that of Iran to which respondents felt somewhat warmer towards after being exposed to the ingroup treatment.[22] While we do see some limited evidence of spillover, the effect of treatment on feelings of warmth toward Turkey was just over *five* times larger than that of the effect of treatment on feelings of warmth towards Iran. Ensuring the treatment is precise is important in general, but is particularly useful here in allowing us to rule out any concerns that our OUTGROUP treatment—which mentions and depicts Iran in a map—is working by simply priming attitudes about "foreign" countries either in general or specifically countries that are associated in some way with either the geographic region of the Middle East or adversaries of the U.S. (e.g., Iran).

Ingroup membership has its reputational privileges

Drawing on SIT, we hypothesized above that ingroup bias would positively influence evaluations of a given actor's "type" (i.e., their reputation). We estimate treatment effects in the manner described above. The resulting treatment effect estimates for reputations are depicted in Figure 8. Several patterns are immediately evident. First, beliefs about traits—attributes or "type"—of Turkey, including *trustworthy*, *tough* and *morally pure*, were all significantly affected by our group categorization treatments, and in the same way. Across the three traits, INGROUP categorization led respondents to perceive Turkey more favorably: as tougher, more honest and more morally righteous compared to the neutral CONTROL group, while our OUTGROUP treatments did not significantly change any beliefs about the target country. These patterns are also evident in respondent beliefs about Turkey's likely behavior, another important aspect of reputation. Across a range of behavioral tendencies—some general like *follow through* and *honor commitments* and others more specific, i.e., *repay debts*, *stand firm in future disputes*,*accept refugees* and *win wars*—ingroup members receive a positive boost. In fact, the only partial exceptions to this pattern are in beliefs about winning wars and standing firm in disputes, where our OUTGROUP treatment significantly shifted respondent beliefs in the negative direction.[23]

Overall, then, H1 finds strong support in our pilot. Our INGROUP treatment positively influenced Turkey's reputation among our U.S.-based respondents. Both aspects of reputational beliefs—traits and predictions of likely behavior—were similarly moved, and this was true for generalized patterns of behavior (e.g., "honor commitments") as well as more specific actions (e.g., "repay debts"). Notably, our study expands the scope of reputation beyond resolve, illustrating the impact of ingroup identity on financial, humanitarian and signaling reputations.

responded to the treatment as we anticipated.

[22]One possible explanation for this is a kind of spillover effect in which given respondents used the information they learned about Turkey to partially inform their views towards Turkey's neighbors.

[23]For our *stand firm* outcome, it was *only* the OUTGROUP treatment which had any impact, while in the *win wars* outcome, the INGROUP treatment pushed beliefs in the opposite (positive) direction.
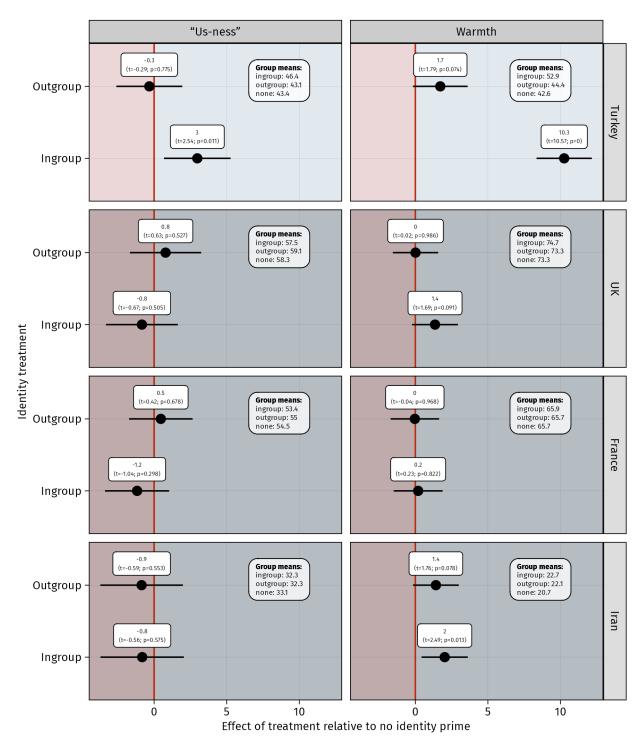
Figure 7: **Group Identity Manipulation Affects Social Categorization of Turkey But (Generally) Not Other States**. Circles represent ATE of treatment relative to control; horizontal bars represent 95% confidence intervals. Dependent variables are all on 0-100 scale with higher values indicating feelings of greater closeness. Results from October 2023 Pilot (N=2,039).
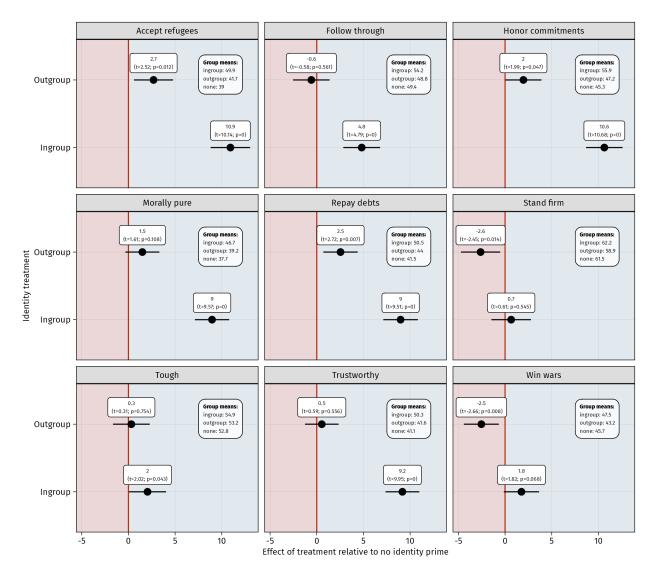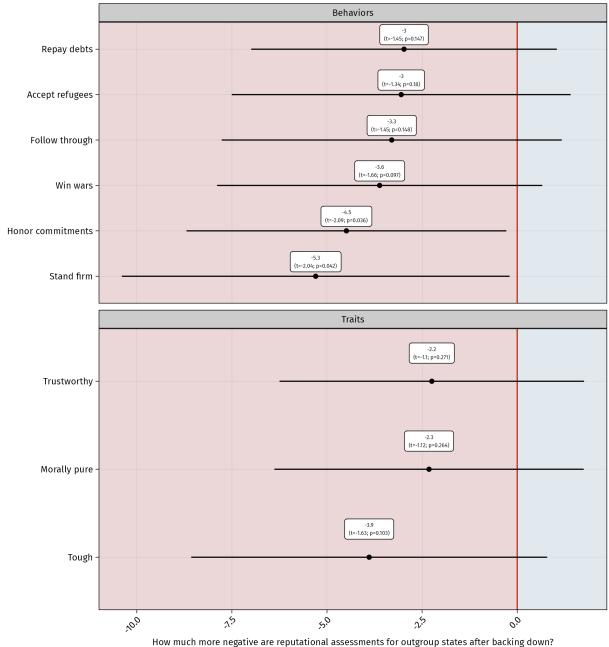
Figure 8: **Ingroup/Outgroup Treatments Affect Reputations**: Circles represent ATE of treatment relative to control; horizontal bars represent 95% confidence intervals. Dependent variables are all on 0-100 scale with higher values indicating more agreement with each of the statements. Results from October 2023 Pilot (N=2,039).

Ingroup Members Are Rewarded Disproportionately for Good Behavior

In the third stage of the experiment, we randomly assigned respondents to arms in which they learn about the past behavior of Turkey in similar international disputes; in particular, they are told that in the past, Turkey either STOOD FIRM or BACKED DOWN in previous disputes. After learning about Turkey's past behavior, we measured our outcome variables for a third time (T_3). Previous work has demonstrated that, in general, past behavior is incorporated into reputations in predictable ways such that standing firm in the past increases observers' estimates of whether that state will stand firm in the future. In our study, we randomized *both* ingroup categorization as well as past behavior, which we now take advantage of to estimate the conditional effect of backing down (relative to standing firm) on reputations across our group categorization treatments. Our estimation strategy differs from above in two ways. First, the dependent variables are now measured at T_3 (just after the past behavior manipulation) instead of T_2 (just after the group identity manipulation). Second, we interact our group identity treatments (INGROUP, OUTGROUP, CONTROL) with our past behavior treatments (STOOD FIRM or BACKED DOWN) in order to produce estimates of the effect of past behavior conditional on the respondent's group identity treatment assignment. Our re-measured outcomes for reputation are the same as above, measuring both traits (e.g., trustworthy) and predicted behavior (e.g., pay back debts).

Our second hypothesis predicted that categorization of a target state as an ingroup member would moderate the impact of observed behavior on reputations, and our results at least partially bear this out. We find evidence suggestive of a dynamic in which the penalty accrued from backing down in the past is somewhat smaller in magnitude for ingroup members than it is for either outgroup members or those in the control arm. We test this hypothesis directly in Figure 9 which plots the estimated difference in the effect of BACKING DOWN between those assigned to INGROUP treatment and those assigned to the OUTGROUP treatment. The quantity presented is an estimate of how much more respondents punish states for backing down when the state in question is characterized as part of the outgroup (relative to an ingroup member). Respondents revise downward their beliefs that Turkey is the kind of state that stands firm (by 5.3 points; $p = .042$) and honors commitments (by 4.5; p=.036) when Turkey assigned to the outgroup treatment compared to those assigned to the ingroup treatment. The same quantity for our other reputational measures is not statistically significant at conventional levels, but always "correctly" signed.

One explanation for the pattern we observe relates to the context of the vignette and the past actions treatment itself. Recall that our "past actions" treatments were exclusively about Turkey's crisis behavior in past disputes, which was also the context of the overall vignette. Whether or not Turkey was described as standing firm in past disputes *should* have an effect (overall) on perceptions related to that domain (e.g., standing firm, tough, honor commitments, follow through) while it would be surprising if such behavior carried over to affect perceptions totally unrelated to whether one had stood firm in past territorial crises (e.g., whether target would accept refugees). Broadly, this is consistent with Dafoe, Renshon and Huth (2014, 386), who argue that past behavior ought to be "more predictive of future behavior to the extent that the contexts are similar."

Figure 9: **Effect of Past Behavior Is (Sometimes) Conditional On Group Categorization**: Here we plot a "diff in diff": the estimated difference in ATE of past behavior (backing down) treatment between those assigned to the INGROUP treatment and those assigned to the OUTGROUP treatment. We learn, e.g., that the "effect of backing down" is more negative (−4.5) for outgroup states (compared to ingroup states). Results from October 2023 Pilot (N=2,039).

Our interpretation—preliminary since the pilot was likely underpowered for these interaction effects—is that both treatments matter here. Past actions treatments overall have effects on a wide range of outcomes, but smaller and less significant effects for some that are less related to crisis behavior, and simultaneously the interaction between past actions and group categorization is also least important for outcomes that we would view as "normatively irrelevant" to crisis behavior, such as repaying debts.[24] In other words, what we find is that group categorization—into either in or outgroups—can have significant effects in how past behavior is consolidated into reputations, *but only (or at least most significantly) for past behavior that is relevant to those particular types of reputations*: standing firm in past disputes does not cause observers to give ingroup members a disproportionate (relative to outgroups) boost in estimations of, e.g., moral purity or willingness to accept refugees. This suggests that reputational spillovers will depend on the behavior, the context in which it is revealed, and how relevant the observer believes the information about an actor's past behavior in one area (conflict) is to the actor's future behavior in another issue area (debt service).

Ingroup Benefits Beyond Reputations

Thus far, we've focused on the question of whether observers perceive target countries differentially based on whether they are categorized as ingroup or outgroup members. Our focus there has been on the direct and immediate consequences for perceptions of the target state as well as how past behavior is differentially incorporated into reputational beliefs based on whether the target is in the ingroup or outgroup. Here, we switch focus to investigate not just how observers *perceive* the target state of Turkey, but their willingness to *act* in cooperating with or helping Turkey in various ways.

We estimate the effect of treatment on our measures of willingness to engage (at T_2) in the same manner as our initial set of analyses. Figure 10 plots the resulting estimates and 95% confidence intervals. The results are strong and consistent with our expectations. On every dimension of helping behavior our INGROUP treatment led to significant boosts in respondents' willingness to help and engage. Interestingly, the effects were largest in the area of military cooperation. Respondents in the INGROUP treatment were 8 points more willing to send U.S. military personnel to help Turkey defend itself and 7.3 points more likely to want to send arms and other materiel than those assigned to the control. This is striking given the fact that the scenario described in the vignette included a description of an ongoing territorial dispute; making the need, but also the costs of supplying such aid clearer. In addition to this increased willingness to help defend Turkey, respondents assigned to the ingroup treatment were also more willing to send Turkey economic aid to help it develop more quickly, to consume Turkish imports, to allow Turkish citizens to live and work in the United States, and to allow more Turkish citizens to tour in the United States.

---

[24]See also Figure 15 in Appendix D.3).
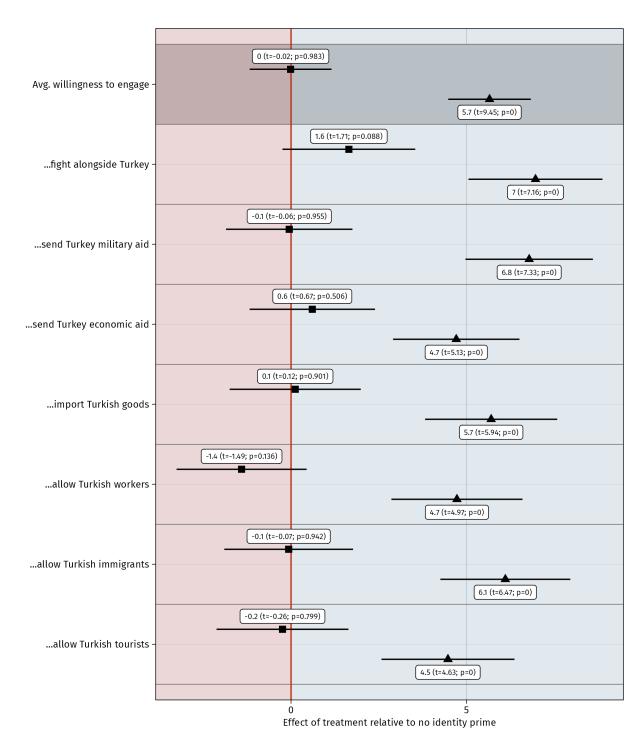
Figure 10: **Benefits to Ingroup Membership in Willingnes to Engage with Turkey**: ■ (Outgroup) and ▲ (Ingroup) represent ATE of treatment relative to control; horizontal bars represent 95% confidence intervals. DVs are all on 0-100 scale with higher values indicating greater willingness to engage in each form of exchange or cooperation. Results from October 2023 Pilot (N= 2,039).

6 Discussion

Recent reviews on reputations in IR have observed that a large gap remains in identifying *whether* reputational inferences differ across states in systematic ways and *what* factors might drive such dynamics (Jervis, Yarhi-Milo and Casler, 2021, 185-86). We argued that one important source of variation in reputational beliefs comes from a factor thus far mostly ignored by reputation scholars: *identity*, and in particular the categorization of other states into either ingroups or outgroups. We sought to make both theoretical and empirical contributions in this paper. Our theory explains how identity dynamics affect reputations in the absence of any behavioral cues as well as the integration of information or observed behavior into reputations. We also used a large literature on helping behavior and ingroup bias to argue that categorization into ingroup/outgroup would affect not just reputations (beliefs) about target states, but observers' willingness to engage with those states in a variety of ways. Finally, our theory of identity and reputation is broad enough to encompass any type of reputation a state may have, advancing our understanding considerably in an area of scholarship dominated by a focus solely on reputations for resolve.

Empirically, we made two signal contributions. First, we fielded a descriptive survey on samples of both U.S. elites and the general public to help address the lacuna in our understanding of the boundaries and content of international identity. From them we learned that elites and the public see international identity in a similar manner, holding nearly identical reputational beliefs about other actors in international affairs as well as overlapping in their views of who is in their ingroup. We also learned that there is a strong relationship between ingroup-ness and valued traits.

Our pilot experiment built on our survey by using a unique design—based around a real country (Turkey)—in which respondents' beliefs about whether the target state was in the ingroup or outgroup were "nudged" in either direction, combined with a repeated measure instrument to allow for more efficient treatment effect estimates. We find broad and strong support for our theory of identity and reputations, overall. First, ingroup categorization has broad and, in some cases, deep effects on baseline reputations. With regard to breadth, we find that our group identity treatments affected a wide range of reputational inferences ranging from perceived magnanimity (e.g., willingness to accept refugees) to perceived resolve (e.g., toughness and willingness to stand firm). Interestingly, the magnitude of these effects varied across several dimensions of reputation. While future work might work to systematically explain this variation, it is striking to us, at least, that the dimensions of reputation *least* affected by our group identity manipulations all relate to the ability of a state to fend for itself in the international system. This would be consistent with the work of Kertzer, Renshon and Yarhi-Milo (2021) in which observers are "intuitive realists" who in some cases rely a great deal on objective markers of power and capability.

Our experiment also shed light on the role that identity plays in incorporating new information into reputational inferences. We show that observers do, in fact, condition how they update based on the how they've categorized the target. Respondents, for example, revised downward their estimates of Turkey's willingness to stand firm and honor its commitments more significantly when assigned to the outgroup treatment (compared to those assigned ingroup member treatment). While

estimates of the analogous quantities for our other dimensions of reputation were all negative, they were not statistically significant. We take this as suggestive evidence that conditional updating occurs generally and will use these results to ensure our full production sample is well-powered to detect these effects of this size if they, in fact, exist.

Finally, our experiment demonstrated that ingroup and outgroup effects go well beyond reputational inferences in IR, extending to preferences for direct engagement. Indeed, we show that those assigned to the ingroup treatment were more willing to cooperate and interact directly with Turkey compared to those assigned to the outgroup treatment or the control. In some cases, the results here are natural implications for our reputational results. Ingroup members are perceived to be tougher, more trustworthy, and more likely to win wars compared to outgroup members so it seems only natural that they would, as we find, also be more willing to help ingroup members defend themselves by sending U.S. military personnel and materiel.[25] These results are important because they suggest that group identity matters in high-stakes, but more infrequent scenarios, while our other engagement results suggest group identity matters as well in more pedestrian (but much more frequent) settings as well. We show that respondents in the ingroup treatment were more willing to host ingroup members as tourists, welcome them as immigrants or guest workers, consume goods imported from ingroup countries, and send economic assistance to help ingroup members develop economically.

Viewing international politics through the lens of social categorization has the potential to address a number of empirical puzzles aside from variation in the updating of reputations. For example, recent work has noted that the Russian invasion of Ukraine led to a host of puzzling changes in attitudes of European citizens. Leveraging a survey in which interviews occurred both pre and post-invasion, Klymak and Vlandas (2022) show that the Russian invasion caused respondents to both value democracy more as well as changed how they saw others by increasing acceptance of immigration from poorer and even non-EU countries. Similar trends were apparent in the U.S., where the Russian invasion led to a large initial spike in favorability of Ukraine among the general public.[26] If it's true that "humans reliably divide the world into *us* and everyone else: *them*" (Cikara and Van Bavel, 2014, 248), then these and other changes in perceptions can be seen as the natural consequences of target states or actors getting accorded the "us treatment."

In the main, then, our results respond to calls to deepen our understanding of how the reputational inferences that observers draw about other actors in world politics might depend on features of not just the actor in question, but the observers themselves. We show that reputations are, indeed, in the "eye of the beholder" (Jervis, 1989). Moreover, we provide suggestive evidence that the effects of these biases are compounding. Identity-based gaps are evident at the baseline and, when presented with new information about an actor, observers, in effect, double down on their biases; revising their estimates of outgroup members more significantly than those of ingroup members.

---

[25]Although it is possible that this result would obtain even in the case where ingroup members were not, in fact, perceived as more likely to win, etc.

[26]https://www.pewresearch.org/global/2023/05/10/americans-hold-positive-feelings-toward-nato-and-ukraine-see-russia-as-an-enemy/.

References

Abdelal, Rawi, Yoshiko M Herrera, Alastair Iain Johnston and Rose McDermott. 2006. "Identity as a Variable." *Perspectives on politics* 4(4):695–711.

Aronow, Peter Michael, Joshua Kalla, Lilla Orr and John Ternovski. 2020. "Evidence of rising rates of inattentiveness on Lucid in 2020." *SocArXiv* pp. 1–10.

Avey, Paul C., Irene Entringer Garcia Blanes, Shauna N. Gillooly, Susan Peterson, Ryan Powers, Andrew Shaver and Michael J. Tierney. 2023. "TRIP Policymaker Survey 2022 Report.".
URL: *https://trip.wm.edu/research/policymaker-surveys/report/Policymaker_Report_2022.pdf*

Avey, Paul C and Michael C Desch. 2014. "What do policymakers want from us? Results of a survey of current and former senior national security decision makers." *International Studies Quarterly* 58(2):227–246.

Avey, Paul C, Michael C Desch, Eric Parajon, Susan Peterson, Ryan Powers and Michael J Tierney. 2021. "Does Social Science Inform Foreign Policy? Evidence from a Survey of US National Security, Trade, and Development Officials." *International Studies Quarterly* 66(1). sqab057.
URL: *https://doi.org/10.1093/isq/sqab057*

Bankert, Alexa, Leonie Huddy and Martin Rosema. 2017. "Measuring partisanship as a social identity in multi-party systems." *Political behavior* 39(1):103–132.

Bankert, Alexa, Ryan Powers and Geoffrey Sheagley. 2022. "Trade Politics at the Checkout Lane: Ethnocentrism and Consumer Preferences." *Political Science Research and Methods* pp. 1–8.

Berinsky, Adam J, Gregory A Huber and Gabriel S Lenz. 2012. "Evaluating online labor markets for experimental research: Amazon. com's Mechanical Turk." *Political analysis* 20(3):351–368.

Berinsky, Adam J, Michele F Margolis, Michael W Sances and Christopher Warshaw. 2021. "Using screeners to measure respondent attention on self-administered surveys: Which items and how many?" *Political Science Research and Methods* 9(2):430–437.

Blair, Graeme, Alexander Coppock and Margaret Moor. 2020. "When to worry about sensitivity bias: A social reference theory and evidence from 30 years of list experiments." *American Political Science Review* 114(4):1297–1315.

Brewer, Marilynn B. 2001. "The many faces of socialidentity: Implications for political psychology." *Political psychology* 22(1):115–125.

Brown, Rupert. 2000. "Social identity theory: Past achievements, current problems and future challenges." *European journal of social psychology* 30(6):745–778.

Brutger, Ryan and Joshua D Kertzer. 2018. "A dispositional theory of reputation costs." *International Organization* 72(3):693–724.

Checkel, Jeffrey T. 1999. "Norms, institutions, and national identity in contemporary Europe." *International studies quarterly* 43(1):83–114.

Chen, Frederick R, Jon CW Pevehouse and Ryan M Powers. 2023. "Great Expectations: the Democratic Advantage in Trade Attitudes." *World Politics* 75(2):316–352.

Cikara, Mina and Jay J Van Bavel. 2014. "The neuroscience of intergroup relations: An integrative review." *Perspectives on Psychological Science* 9(3):245–274.

Clare, Joe and Vesna Danilovic. 2010. "Multiple audiences and reputation building in international conflicts." *Journal of Conflict Resolution* 54(6):860–882.

Clifford, Scott, Geoffrey Sheagley and Spencer Piston. 2021. "Increasing precision without altering treatment effects: Repeated measures designs in survey experiments." *American Political Science Review* 115(3):1048–1065.

Coppock, Alexander and Oliver A McClellan. 2019. "Validating the demographic, political, psychological, and experimental results obtained from a new source of online survey respondents." *Research & Politics* 6(1):2053168018822174.

Crescenzi, Mark JC. 2018. *Of Friends and Foes: Reputation and Learning in International Politics.* Oxford University Press.

Cuhadar, Esra and Bruce Dayton. 2011. "The social psychology of identity and inter-group conflict: From theory to practice." *International Studies Perspectives* 12(3):273–293.

Curley, Tyler M. 2009. "Social identity theory and EU expansion." *International Studies Quarterly* 53(3):649–668.

Dafoe, Allan and Devin Caughey. 2016. "Honor and war: Southern US presidents and the effects of concern for reputation." *World politics* 68(2):341–381.

Dafoe, Allan, Jonathan Renshon and Paul Huth. 2014. "Reputation and status as motives for war." *Annual Review of Political Science* 17:371–393.

Deaux, Kay, Anne Reid, Kim Mizrahi and Kathleen A Ethier. 1995. "Parameters of social identity." *Journal of personality and social psychology* 68(2):280.

Distelhorst, Greg and Yue Hou. 2014. "Ingroup bias in official behavior: A national field experiment in China." *Quarterly Journal of Political Science* 9(2):203–230.

Druckman, Daniel. 1994. "Nationalism, patriotism, and group loyalty: A social psychological perspective." *Mershon international studies review* 38(Supplement_1):43–68.

Fernández-Reino, Mariña, Valentina Di Stasio and Susanne Veit. 2023. "Discrimination unveiled: a field experiment on the barriers faced by Muslim women in Germany, the Netherlands, and Spain." *European Sociological Review* 39(3):479–497.

Findley, Michael G, Daniel L Nielson and Jason C Sharman. 2013. "Using Field Experiments in International Relations: A Randomized Study of Anonymous Incorporation1." *International Organization* 67(4):657–693.

Greene, Steven. 1999. "Understanding party identification: A social identity approach." *Political psychology* 20(2):393–403.

Gries, Peter Hays. 2005. "Social Psychology and the Identity-Conflict Debate: Is a 'China Threat'Inevitable?" *European Journal of International Relations* 11(2):235–265.

Holsti, Ole R. 1967. "Cognitive dynamics and images of the enemy." *Journal of International Affairs* 21(1):16–39.

Huddy, Leonie. 2001. "From social to political identity: A critical examination of social identity theory." *Political psychology* 22(1):127–156.

Huddy, Leonie, Lilliana Mason and Lene Aarøe. 2015. "Expressive partisanship: Campaign involvement, political emotion, and partisan identity." *American Political Science Review* 109(1):1–17.

Huff, Connor and Dustin Tingley. 2015. ""Who are these people?" Evaluating the demographic characteristics and political preferences of MTurk survey respondents." *Research & Politics* 2(3):2053168015604648.

Hymans, Jacques EC. 2002. Applying social identity theory to the study of international politics: a caution and an agenda. In *annual meeting of the International Studies Association, New Orleans*. pp. 24–27.

Islam, Gazi. 2014. "Social identity theory." *Journal of personality and Social Psychology* 67:741–763.

Jervis, Robert. 1989. "Additional thoughts on political psychology and rational choice." *Political Psychology* 10(3):511–515.

Jervis, Robert, Keren Yarhi-Milo and Don Casler. 2021. "Redefining the Debate over Reputation and Credibility in International Security: Promises and Limits of New Scholarship." *World Politics* 73(1):167–203.

Johnson, Dominic and Dominic Tierney. 2003. "Essence of victory: Winning and losing international crises." *Security Studies* 13(2):350–381.

Johnson, Dominic DP, Dominic Tierney et al. 2006. *Failing to win: Perceptions of victory and defeat in international politics.* Harvard University Press.

Katzenstein, Peter J. 1996. *The culture of national security: Norms and identity in world politics.* Columbia University Press.

Kertzer, Joshua D. 2016. *Resolve in international politics*. Princeton University Press.

Kertzer, Joshua D. 2022. "Re-assessing elite-public gaps in political behavior." *American Journal of Political Science* 66(3):539–553.

Kertzer, Joshua D and Dustin Tingley. 2018. "Political psychology in international relations: beyond the paradigms." *Annual Review of Political Science* 21(1):319–339.

Kertzer, Joshua D and Jonathan Renshon. 2022. "Experiments and surveys on political elites." *Annual Review of Political Science* 25:529–550.

Kertzer, Joshua D, Jonathan Renshon and Keren Yarhi-Milo. 2021. "How Do Observers Assess Resolve?" *British Journal of Political Science* 51(1):308–330.

Kertzer, Joshua D, Jonathan Renshon and Keren Yarhi-Milo. 2024. "Reputations in the Wild: Reputation Costs from the Syrian and Crimean Crises Among Leaders and Citizens.".

Kertzer, Joshua D, Marcus Holmes, Brad L LeVeck and Carly Wayne. 2022. "Hawkish Biases and Group Decision Making." *International Organization* pp. 1–36.

Klymak, Margaryta and Tim Vlandas. 2022. "The Impact of Russia's Invasion on European Attitudes." *Available at SSRN 4315398* .

Kowert, Paul and Jeffrey Legro. 1996. The culture of national security: Norms and identity in world politics. In *Norms, identity, and their limits: A theoretical reprise.* pp. 451–97.

Kydd, Andrew H. 2007. *Trust and mistrust in international relations*. Princeton University Press.

Larson, Deborah Welch and Alexei Shevchenko. 2003. "Shortcut to greatness: The new thinking and the revolution in Soviet foreign policy." *International Organization* 57(1):77–109.

Levine, Mark, Amy Prosser, David Evans and Stephen Reicher. 2005. "Identity and emergency intervention: How social group membership and inclusiveness of group boundaries shape helping behavior." *Personality and social psychology bulletin* 31(4):443–453.

Lo, Adeline, Jonathan Renshon and Lotem Bassan-Nygate. 2023. "A Practical Guide to Dealing with Attrition in Political Science Experiments." *Journal of Experimental Political Science* pp. 1–15.

Löwenheim, Oded. 2003. ""Do Ourselves Credit and Render a Lasting Service to Mankind": British Moral Prestige, Humanitarian Intervention, and the Barbary Pirates." *International Studies Quarterly* 47(1):23–48.

Lundberg, Ian, Rebecca Johnson and Brandon M Stewart. 2021. "What is your estimand? Defining the target quantity connects statistical evidence to theory." *American Sociological Review* 86(3):532–565.

Lupton, Danielle L. 2020. *Reputation for Resolve.* Cornell University Press.

Mason, Lilliana, Julie Wronski and John V Kane. 2021. "Activating animus: The uniquely social roots of Trump support." *American Political Science Review* 115(4):1508–1516.

McGuire, William J, Claire V McGuire, Pamela Child and Terry Fujioka. 1978. "Salience of ethnicity in the spontaneous self-concept as a function of one's ethnic distinctiveness in the social environment." *Journal of personality and social psychology* 36(5):511.

Mendeloff, David. 2008. "Pernicious history'as a cause of national misperceptions: Russia and the 1999 Kosovo War." *Cooperation and Conflict* 43(1):31–56.

Mercer, Jonathan. 1995. "Anarchy and identity." *International Organization* 49(2):229–252.

Mercer, Jonathan. 1997. "Reputation and rational deterrence theory." *Security Studies* 7(1):100–113.

Mercer, Jonathan. 2005. "Rationality and psychology in international politics." *International organization* 59(1):77–106.

Mercer, Jonathan. 2010. *Reputation and International Politics.* Cornell University Press.

Mercer, Jonathan. 2023. "Racism, Stereotypes, and War." *International Security* 48(2):7–48.

Murray, Michelle. 2018. *The struggle for recognition in international relations: status, revisionism, and rising powers.* Oxford University Press.

Otten, Sabine. 2016. "The Minimal Group Paradigm and its maximal impact in research on social categorization." *Current Opinion in Psychology* 11:85–89.

Pager, Devah. 2003. "The mark of a criminal record." *American journal of sociology* 108(5):937–975.

Pinto, Diego Costa, Adilson Borges, Márcia Maurer Herter and Mário Boto Ferreira. 2020. "Reducing ingroup bias in ethical consumption: The role of construal levels and social goodwill." *Business Ethics Quarterly* 30(1):31–63.

Politi, Emanuele, Jessica Gale, Antoine Roblain, Magdalena Bobowik and Eva GT Green. 2023. "Who is willing to help Ukrainian refugees and why? The role of individual prosocial dispositions and superordinate European identity." *Journal of Community & Applied Social Psychology* 33(4):940–953.

Powers, Ryan. 2023. "Is Context Pretext? Domestic Audiences and the Situational Politics of Noncompliance.".
URL: *https://ryanpowers.net/files/ContextPretextPowers.pdf*

Reicher, Stephen. 2004. "The Context of Social Identity: Domination, Resistance, and Change." *Political Psychology* pp. 921–945.

Renshon, Jonathan, Allan Dafoe and Paul Huth. 2018. "Leader influence and reputation formation in world politics." *American Journal of Political Science* 62(2):325–339.

Rousseau, David L and Rocio Garcia-Retamero. 2007. "Identity, power, and threat perception: A cross-national experimental study." *Journal of Conflict Resolution* 51(5):744–771.

Sambanis, Nicholas, Stergios Skaperdas and William C Wohlforth. 2015. "Nation-building through war." *American Political Science Review* 109(2):279–296.

Schelling Thomas, C. 1966. "Arms and influence.".

Sinclair, Samantha, Mark Granberg and Towe Nilsson. 2023. "Love thy (Ukrainian) neighbour: Willingness to help refugees depends on their origin and is mediated by perceptions of similarity and threat." *British Journal of Social Psychology* .

Smith, Rogers M. 2004. "Identities, interests, and the future of political science." *Perspectives on Politics* 2(2):301–312.

Sniderman, Paul M, Louk Hagendoorn and Markus Prior. 2004. "Predisposing factors and situational triggers: Exclusionary reactions to immigrant minorities." *American political science review* 98(1):35–49.

Spears, Russell. 2021. "Social influence and group identity." *Annual review of psychology* 72:367–390.

Stagnaro, Michael Nicholas, James Druckman, Antonio Alonso Arechar, Robb Willer, David Rand et al. 2024. "Representativeness Versus Attentiveness: A Comparison Across Nine Online Survey Samples.".

Streich, Philip and Jack S Levy. 2016. "Information, commitment, and the russo-japanese war of 1904–1905." *Foreign Policy Analysis* 12(4):489–511.

Ternovski, John and Lilla Orr. 2022. "A Note on Increases in Inattentive Online Survey-Takers Since 2020." *Journal of Quantitative Description: Digital Media* 2.

Tetlock, Philip E and James M Goldgeier. 2000. "Human nature and world politics: Cognition, identity, and influence." *International Journal of Psychology* 35(2):87–96.

Tomz, Michael. 2007. *Reputation and International Cooperation*. Princeton University Press.

Walter, Barbara F. 2009. *Reputation and civil war: Why separatist conflicts are so violent*. Cambridge University Press.

Weisiger, Alex and Keren Yarhi-Milo. 2015. "Revisiting reputation: How past actions matter in international politics." *International Organization* 69(2):473–495.

West, Emily A and Shanto Iyengar. 2020. "Partisanship as a social identity: Implications for polarization." *Political Behavior* pp. 1–32.

White, Katherine and Darren W Dahl. 2007. "Are all out-groups created equal? Consumer identity and dissociative influence." *Journal of Consumer Research* 34(4):525–536.

Wohlforth, William C. 2009. "Unipolarity, status competition, and great power war." *World politics* 61(1):28–57.

Yarhi-Milo, Keren. 2018. *Who Fights for Reputation.* Princeton University Press.

Appendix Materials for "Identity and Reputation in World Politics"

Appendix

Table of Contents

A   Initial Descriptive Survey

A.1   Demographic Distribution for Public Sample

	Lucid (N = 1,599)	2020 CCES
Gender		
Male	797 (49.84%)	48.5%
Female	802 (50.16%)	51.5%
Income		
Less than $30,000	647 (40.46%)	26.1%
Between $30,000 and $59,999	428 (26.77%)	25.41%
Between $60,000 and $149,999	420 (26.27%)	30.8%
$150,000 or more	82 (5.13%)	6.61%
Prefer not to say	22 (1.38%)	11.08%
Age		
18–29	386 (24.14%)	21.04%
30–39	328 (20.51%)	15.93%
40–49	268 (16.76%)	13.73%
50–59	241 (15.07%)	17.56%
60–69	234 (14.63%)	17.94%
70+	142 (8.88%)	13.8%
Region		
Northeast	325 (20.33%)	18%
Midwest	297 (18.57%)	21.62%
South and Central	599 (37.46%)	40.37%
West	378 (23.64%)	20.01%
Party ID		
Democrat	595 (37.21%)	42.2%
Republican	355 (22.20%)	37.95%
Independent	494 (30.89%)	14.35%
Other	155 (9.69%)	5.5%
Education		
Some high school or less	79 (4.94%)	8.85%
High school graduate	435 (27.20%)	28.56%
Some college	344 (21.51%)	21.36%
2 year degree	115 (7.19%)	10.19%
4 year degree	364 (22.76%)	19.67%
Post-grad	255 (15.95%)	11.37%
Other/Prefer not to say	7 (0.44%)	0%
Ethnicity		
White	1,038 (64.92%)	70.28%
Hispanic	87 (5.44%)	10.08%
Black	221 (13.82%)	13.15%
American Indian or Alaska Native	22 (1.38%)	0.71%
Asian	79 (4.94%)	4.38%
Other	141 (8.82%)	1.41%
Prefer not to say	11 (0.69%)	0%

A.2 Measuring Identities in International Politics

1. Attention Check 1

– "In the grid below, you will see a series of statements. Please tell us whether you agree or disagree with each statement." A "grid" question with eight items (from agree strongly to disagree strongly), embedded in which is the statement "two is greater than one."

2. Group and Describe

(a) "In discussing world affairs, people sometimes use words like 'us' and 'them': 'us' to describe our country and other countries that are like us in some way, or part of our group, and 'them' to describe ones that are not like us and not part of the same group. **We're interested in your views on which countries are in the 'us' group and which are in the 'them' group in international politics.** Below, you will see a list of countries. Please categorize as either being in the 'us' group or 'them' group."

– Respondents see a list of 10 countries randomly selected from the top 25 most powerful countries as measured by the Correlates of War "composite indicator of national capabilities" measure.

(b) "Above, you placed countries in the '**them**' group. In the spaces below, please list some words or phrases that describe what you think these countries have in common. What makes them part of the '**them**' group?" Word listing entry for respondents.

(c) "Above, you placed countries in the '**us**' group. In the spaces below, please list some words or phrases that describe what you think these countries have in common. What makes them part of the '**us**' group?" Word listing entry for respondents.

3. Attention Check 2

– "There are many important issues facing our country today. Research shows that issues people think are important can affect their views on other issues. We also want to know if you are paying attention. Please ignore the question and put 'crime' in the top position and 'unemployment' in the bottom position. Leave the rest of the issues in the same order. Please rank the following issues facing the nation from 1 (most important) to 7 (least important). You can change your rankings by dragging and dropping different issues."

4. Rate Countries and Groups

– (Countries) "We're interested in perceptions of other members of the global community. **Below, we ask for your level of agreement with a number of statements about countries and groups of countries around the world.** Please give your impressions of the country or group, whether or not you know much about them. The country of [insert country name from prior question] is..."

– Respondents rate items (trustworthy, likely to stand firm against enemies, morally pure, likely to honor commitments, likely to pay its debts, similar to the United States, powerful) from 0 (no agreement at all) to 100 (total agreement)

– Subjects rate 3 countries chosen at random from the 10 countries they grouped in previous question.

– (Groups) "We're interested in perceptions of other members of the global community. **Below, we ask for your level of agreement with a number of statements about countries and groups of countries around the world.** Please give your impressions of the country or group, whether or not you know much about them. Countries that can be described as [group name] are..."

– Groups are drawn at random from: "open and democratic" , "closed and authoritarian", "capitalist", "socialist", "communist", "members of the United Nations Security Council", "Western", "great powers", "wealthy", "developing," The dimensions on which they are ranked are same as above. Subjects rate 3 groups.

5. Attention Check 3

   – "In the grid below, you will see a series of statements. Please tell us whether you agree or disagree with each statement." 8 statements, embedded in which is the statement "World War I came after World War II."

6. Closed-Choice Identity Question: subjects randomized into either (a) or (b) version below.

   (a) "People can define or describe themselves in terms of social categories or groups to which they belong. In a similar way, we can think of **countries** as belonging to different groups or categories or having different identities.

   Below are a list of group memberships and identities that may be relevant to how you think about the United States in the world. **For each group, please rank its importance, so that items ranked 0 are groups that the United States is not part of or are not very important to how you think of the U.S. in the world. Items ranked 100 are terms that are very important to how you think about the United States in the world.**"

   – Groups: North-American, Western Hemisphere, Western Civilization, Free, Democratic, Superpower, Capitalist, English-Speaking, Member of the United Nations, Permanent Member of the UN Security Council, Predominantly Caucasian, Ethnic Melting Pot.

   (b) "In discussing world affairs, people sometimes use words like '**us**' and '**them**': 'us' to describe our country and other countries that are like us in some way, or part of are group, and 'them' to describe ones that are not like us and not part of the same group. We're interested in your views on what things define 'us' in international politics.

   For each item below, please rank its importance, so that items ranked closer to 0 are groups that the United States is not part of and/or terms that don't really describe the United States, while items ranked closer to 100 are terms that describe the United States very well and/or are groups that the United States is part of."

   – Same groups as above.

7. Attention Check 4

   – "We are also interested in what sections people like to read in the newspaper. This might affect what they learn from articles and how they feel about the issues discussed in them. We also want to see if people are reading the questions carefully. To show that you've read this much, please mark both the classified and none of the above boxes below. That's right, just select these two options only. Regardless of how frequently you read the newspaper, what would you say are your favorite newspaper sections to read? (please check all that apply)."

8. Open-Ended Identity Question: subjects randomized into either (a) or (b) version below.

   (a) "People can define or describe themselves in terms of social categories or groups to which they belong. In a similar way, we can think of **countries** as belonging to different groups or categories or having different identities.

For example, some group memberships, or identities, that might be relevant for the United States are: 'English-speaking' or 'democracy.' **What are some groups or social categories that are relevant to you as you think about the United States and its identity in the world?** Please list as many as you can think of."

(b) "In discussing world affairs, people sometimes use words like '**us**' and '**them**': "us" to describe our country and other countries that are like us in some way, or part of our group, and "them" to describe ones that are not like us and not part of the same group. We're interested in your views on what things define 'us' in international politics. For example, some may label democracies as part of the "us" group and dictatorships part of the 'them' group. **What are groups or social categories that you would say are in the 'us' group?**"

9. Constant-Sum Identity Tradeoff

– "Think again about the identities that are important to the United States that we asked about above.

Use the boxes to the right to indicate how important each of the following categories are to how you think about our country's place and identity in the international community.

Lower values imply less importance. You have a total of 100 'points' to spend across all categories. You may allocate them across categories as you see fit. Those that are of no importance can take a value of zero."

– Groups: North-American, Western Hemisphere, Western Civilization, Free, Democratic, Superpower, Capitalist, English-Speaking, Member of the United Nations, Permanent Member of the UN Security Council, Predominantly Caucasian, Ethnic Melting Pot.

A.3 Three lessons from our descriptive survey

In the main text, we discuss three main lessons from our descriptive surveys. We present figures based on the results of our elite and public samples. Here we show that results are similar if we also include IR scholars.

Elites and the public hold similar reputational beliefs about other states

B Pilot Description

This study proceeded as follows:

1. Measure baseline beliefs about Turkey's reputation, willingness to engage, warmth, and "us-ness" (T_1):

(a) Reputation: The country of Turkey is... (Response: 0 = No agreement; 100 = Total agreement)

   i. ...likely to stand firm against its enemies

   ii. ...tough

Figure 11: **Elites and the public agree on who is in the ingroup.** We asked respondents to categorize countries as either "us" or "them." The scatter plot shows that countries that elites were most likely to categorize as "us" were also categorized as "us" by the public.

Figure 12: **Elites and the public hold similar reputational beliefs.** We asked respondents to indicate their level of agreement with a variety of statements about different countries.

Figure 13: **Positive traits are correlated with "us-ness."**.

iii. ...likely to honor its international commitments

iv. ...trustworthy

v. ...morally pure

vi. ...likely to win a war if it chooses to start one

vii. ...likely to do the things it says it will do

viii. ...likely to repay its debts

ix. ...likely to accept refugees from other countries

(b) Willingness to Engage: International relations sometimes requires working with or even defending other members of the global community. How willing would you to do the following with or for the country of Turkey? (Response: 0 = No agreement; 100 = Total agreement)
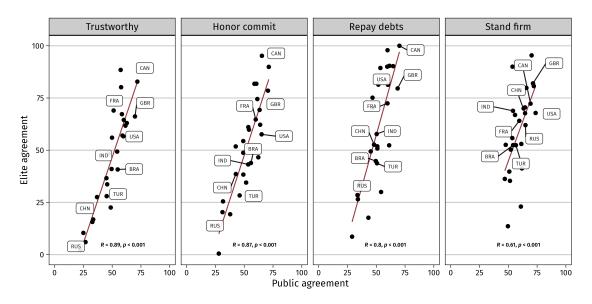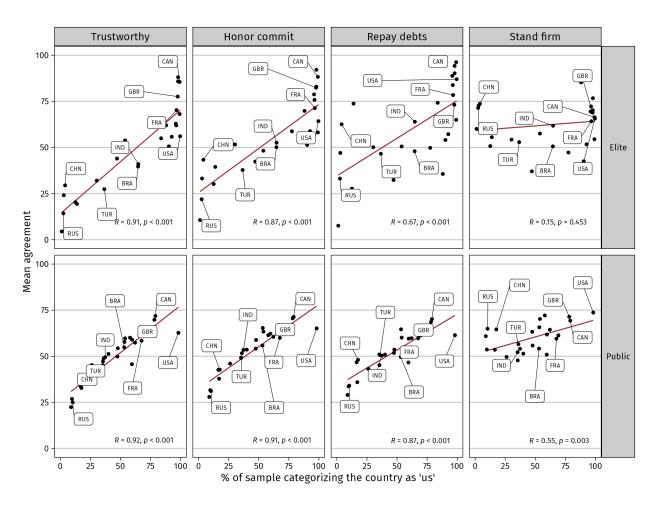
i. Allow consumer goods from Turkey to be sold in the United States

ii. Allow tourists from Turkey into the United States

iii. Allow Turkish citizens to live and work in the United States

iv. Allow Turkish citizens to become U.S. citizens

v. Send money to Turkey to help it grow economically

vi. Send guns, missiles, and other military supplies to help Turkey defend itself against an attack by another country

vii. Send U.S. military personnel to help defend Turkey against an attack by another country

(c) Us vs. Them: In discussing world affairs, people sometimes use words like "us" and "them": "us" to describe our country and other countries that are like us in some way, or part of our group, and "them" to describe ones that are not like us and not part of the same group. Where would you place the following countries along an "us" to "them" continuum? (Response: 0 = Entirely us; 100 = Entirely them. Reverse coded for analysis).

i. Turkey

ii. France

iii. Iran

iv. United Kingdom

(d) Warmth: We are interested in your feelings toward other members of the global community. We would like you to indicate how warmly you feel toward each of the countries listed below. (Response: 0 = very cold/unfriendly; 100 = very warm/friendly).

i. Turkey

ii. France

iii. Iran

    iv. United Kingdom

2. All respondents read a baseline scenario about what is described as the beginnings of a diplomatic conflict between two countries, Turkey, and another country, B:

> Turkey is currently involved in a public territorial dispute with another country. Turkey maintains that the territory belongs to Turkey, but the other country has disputed this claim. Because of the risk that disputes of this nature might escalate, there has been significant media coverage of the confrontation in Turkey and around the world.

3. Subjects randomized to receive an identity prime of "no additional info," "ingroup prime" or "outgroup prime" ($p=1/3$ each). The details of these manipulations are given in the main text.

4. Measure beliefs about Turkey's reputation (T_2).

5. All subjects randomized into further information about Turkey's behavior:

   - Past behavior: no information, stood firm in the past, backed down in the past

6. All subjects complete post-vignette DV measurement (T_3)
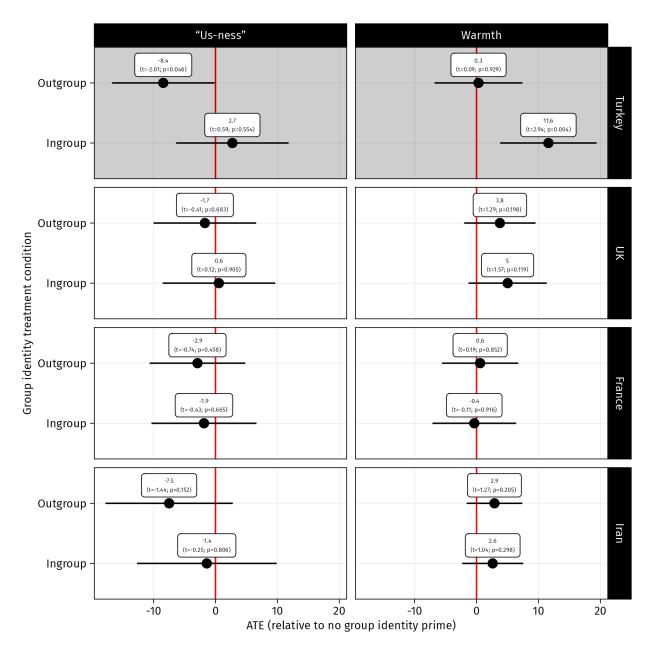
C Pre-Pilot Results

Figure 14: Results from survey of October Pre-Pilot of Treatment (N=200).

D Pilot Results

D.1 Demographic Distribution

	Distribution (N = 2,158)
Gender	
Male	1,044 (48.38%)
Female	1,114 (51.62%)
Prefer not to answer	0 (0.00%)
Income	
Less than $49,999	1,090 (50.51%)
$50,000 to $99,999	716 (33.18%)
$100,000 to $149,999	220 (10.19%)
$150,000 or more	125 (5.79%)
Prefer not to answer	7 (0.32%)
Age	
18–29	365 (16.91%)
30–39	425 (19.69%)
40–49	414 (19.18%)
50–59	379 (17.56%)
60–69	335 (15.52%)
70+	240 (11.12%)
Region	
Northeast	387 (17.93%)
Midwest	437 (20.25%)
South and Central	849 (39.34%)
West	485 (22.47%)
Party ID	
Democratic	995 (46.11%)
Independent/Other	395 (18.30%)
Republican	768 (35.59%)
Prefer not to say	0 (0.00%)
Education	
Some high school or less	57 (2.64%)
High school graduate	396 (18.35%)
Some college or vocational training	618 (28.64%)
2 year degree	251 (11.63%)
4 year degree	522 (24.19%)
Post-grad	313 (14.50%)
Prefer not to say	1 (0.05%)
Ethnicity	
White	1,581 (73.26%)
Black	229 (10.61%)
American Indian or Alaska Native	42 (1.95%)
Asian	89 (4.12%)
Prefer not to say	0 (0.00%)

D.2 Attention Checks

We ask several attention checks in our study. The first asked respondents to select their favorite color from a list but included special instructions in a preamble to ignore the question and select "red" and "green." In all, about 32 percent (2202/6846) of those who agreed to participate answered this question correctly. We embedded a second attention check in a grid question related to the respondent's foreign policy dispositions. The question told the respondent to select "disagree" in a given row of the matrix. About 92 percent of those who passed the first attention check passed our second attention check. Those respondents who failed either attention check (70.2%) were re-directed back to the sample provider; they did not complete any further questions or participate in the experiments.

We screen for inattention because Lucid has acquired a reputation for inattentive respondents (Aronow et al., 2020). Indeed, Stagnaro et al. (2024) anticipated that their Lucid sample would yield "failed attention rates of between 1/3 - 2/3" in their recent comparative study of online sample providers. Two-thirds is not so far off our 70 percent failure rate and our attention checks were somewhat more demanding than their "CAPTCHA-style" attention checks (e.g., their first attention check was: "Please write 'twenty-five' using numbers. (free response box)"). While Lucid samples lag the field somewhat in attentiveness, they lead the field in representativeness. Stagnaro et al. (2024) compared samples from Lucid, Prolific, mTurk and others and concluded that Lucid and another provider, Bovtiz, provided samples that were "the closest to representative probability samples on demographic and attitudinal representativeness." While their study showed that filtering inattentive respondents decreases representativeness somewhat, even the most aggressive filters yielded a sample that was closer to their national probability sample benchmark than all other providers except for Bovitz.

Outcome	n	percent
(pre-treatment): Select 'Green' and 'Red' from multiple select		
Failed	4662	66.5%
Passed	2350	33.5%
Total	7012	-
(pre-treatment): Select 'Disagree' from matrix question		
Failed	182	7.7%
Passed	2168	92.3%
Total	2350	-
(post-treatment): Select 'Neither Agree nor Disagree' from matrix question		
Failed	26	1.3%
Passed	2008	98.7%
Total	2034	-

Table 5: Distribution of attention check outcomes. Pre-treatment questions used to select for attentive respondents. Post-treatment question used to assess how durable attention levels were over the course of the experiment.

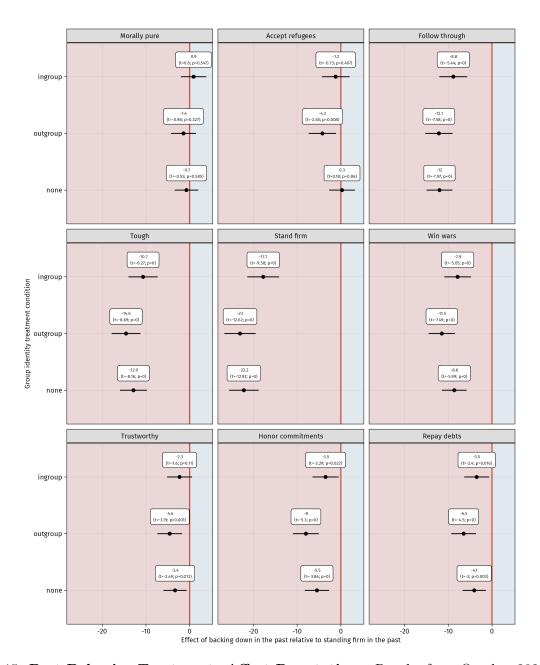## D.3 Reputation Outcomes Conditional on Past Actions

Figure 15: **Past Behavior Treatments Affect Reputations**: Results from October 2023 Pilot (N=2,039).

D.4   Results Tables for Models Used to Produce Figures

	Warmth				Us v. Them			
	Turkey	France	Iran	UK	Turkey	France	Iran	UK
Identity Manipulation (Baseline: Control)								
Outgroup	1.720+	-0.034	1.418+	0.014	-0.330	0.463	-0.861	0.791
	(0.961)	(0.851)	(0.804)	(0.796)	(1.158)	(1.116)	(1.451)	(1.249)
Ingroup	10.254***	0.193	2.022*	1.357+	2.975*	-1.172	-0.821	-0.841
	(0.970)	(0.859)	(0.812)	(0.804)	(1.169)	(1.126)	(1.465)	(1.261)
Num.Obs.	2054	2054	2054	2054	2054	2054	2054	2054
R2	0.535	0.601	0.636	0.651	0.342	0.450	0.422	0.460

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6: Results used to produce Figure 7 Models estimated via OLS. All dependent variables are on a 0–100 scale. All models control for pre-treatment (T_1) measures of the DVs (coefficients omitted).

	Stand Firm	Tough	Honor Commit	Trust
Identity Manipulation (Baseline: Control)				
Outgroup	-2.62*	0.31	1.97*	0.54
	(1.07)	(1.00)	(0.99)	(0.92)
Ingroup	0.65	2.05*	10.64***	9.18***
	(1.08)	(1.01)	(1.00)	(0.92)
Num.Obs.	2054	2054	2054	2054
R2	0.399	0.495	0.513	0.572

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 7: Results used to produce, in part, Figure 8. Models estimated via OLS. All dependent variables are on a 0–100 scale. All models control for pre-treatment (T_1) measures of the DVs (coefficients omitted).

	Pure	Win	Follow thru	Repay debts	Accept refugees
Identity Manipulation (Baseline: Control)					
Outgroup	1.49	-2.54**	-0.58	2.55**	2.69*
	(0.93)	(0.96)	(1.00)	(0.94)	(1.07)
Ingroup	8.96***	1.76+	4.82***	8.98***	10.93***
	(0.94)	(0.96)	(1.01)	(0.95)	(1.08)
Num.Obs.	2054	2054	2054	2054	2054
R2	0.546	0.485	0.492	0.555	0.453

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Table 8: Results used to produce, in part, Figure 8. Models estimated via OLS. All dependent variables are on a 0–100 scale. All models control for pre-treatment (T_1) measures of the DVs (coefficients omitted).

	Stand Firm	Tough	Honor Commit	Trust
Identity Manipulation (Baseline: Ingroup)				
Control	0.49	-2.60	-7.47***	-7.74***
	(1.76)	(1.62)	(1.45)	(1.39)
Outgroup	-0.38	-1.20	-4.47**	-7.03***
	(1.84)	(1.69)	(1.52)	(1.45)
Past Behavior Manipulation (Baseline: Stand Firm)				
Back down	-17.74***	-10.66***	-3.49*	-2.32
	(1.85)	(1.70)	(1.53)	(1.46)
Interactions				
Control * Back down	-4.42+	-2.19	-1.98	-1.03
	(2.52)	(2.32)	(2.08)	(1.98)
Outgroup * Back down	-5.30*	-3.89	-4.49*	-2.25
	(2.60)	(2.39)	(2.14)	(2.04)
Num.Obs.	2038	2038	2038	2038
R2	0.375	0.427	0.496	0.533

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Table 9: Results used to produce Figure 15. Models estimated via OLS. All dependent variables are on a 0–100 scale. All models control for pre-treatment (T_1) measures of the DVs (coefficients omitted).

	Pure	Win	Follow thru	Repay debts	Accept refugees
Identity Manipulation (Baseline: Ingroup)					
Control	-7.38***	-1.35	-2.99+	-7.29***	-8.15***
	(1.41)	(1.48)	(1.55)	(1.39)	(1.54)
Outgroup	-5.48***	-1.16	-2.99+	-5.09***	-3.59*
	(1.47)	(1.55)	(1.62)	(1.45)	(1.61)
Past Behavior Manipulation (Baseline: Stand Firm)					
Back down	0.89	-7.85***	-8.84***	-3.51*	-1.18
	(1.48)	(1.55)	(1.62)	(1.46)	(1.62)
Interactions					
Control * Back down	-1.64	-0.76	-3.16	-0.55	1.44
	(2.01)	(2.12)	(2.21)	(1.99)	(2.21)
Outgroup * Back down	-2.32	-3.62+	-3.30	-2.97	-3.05
	(2.07)	(2.18)	(2.28)	(2.05)	(2.27)
Num.Obs.	2038	2038	2038	2038	2038
R2	0.521	0.420	0.446	0.530	0.462

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Table 10: Results used to produce Figure 9. Models estimated via OLS. All dependent variables are on a 0–100 scale. All models control for pre-treatment (T_1) measures of the DVs (coefficients omitted).

	Mean	Goods	Tourists	Immigrants	Workers	Econ Aid	Mil Aid	Fight For
Identity Manipulation (Baseline: Control)								
Outgroup	-0.068	-0.027	-0.384	-0.503	-1.017	0.225	0.179	1.052
	(0.594)	(0.914)	(0.907)	(0.894)	(0.894)	(0.861)	(0.896)	(0.926)
Ingroup	5.705***	5.664***	4.923***	5.851***	4.869***	4.955***	6.954***	6.718***
	(0.598)	(0.920)	(0.914)	(0.900)	(0.900)	(0.867)	(0.902)	(0.933)
Num.Obs.	2054	2054	2054	2054	2054	2054	2054	2054
R2	0.794	0.687	0.704	0.739	0.733	0.672	0.672	0.628

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 11: Results used to produce Figure 10. Models estimated via OLS. All dependent variables are on a 0–100 scale. All models control for pre-treatment (T_1) measures of the DVs (coefficients omitted).