Breaking Bad? How Survey Experiments Prime Americans for War Crimes

Charli Carpenter, Alexander H. Montgomery, and Alexandria Nylen

What affects Americans' sensitivity to international laws and norms on the use of force? A wealth of recent IR literature tackles this question through experimental surveys using fictional scenarios and treatments to explore precisely when Americans would approve of government policies that would violate the laws of war. We test whether such survey experiments may themselves be affecting public sensitivity to these norms—or even Americans' understanding of the content of the norms themselves. We show that being invited to express a preference regarding war crimes in survey settings has a negative impact on Americans' understanding of US legal and ethical obligations in war and that reporting previous findings can inflate support for war crimes. We conclude with suggestions for future experimental survey design in international relations and international law.

A cottage industry has emerged in IR aiming to measure citizen attitudes toward the laws of war (Chilton 2015; Chilton and Tingley 2013; Rathbun and Stein 2019). Many recent studies tackle this question through experimental surveys that use fictional scenarios and treatments to explore precisely *when* citizens would support bombing civilians, engaging in preventive war, outsourcing kill decisions to machines, torturing detainees, or conducting extrajudicial executions with drones (Haworth, Sagan, and Valentino 2019; Horowitz 2016; Kreps and Wallace 2016; Press, Sagan, and Valentino 2013; Walsh 2015). The vignettes in question generally present respondents with a scenario in which it might be tempting to commit a violation of international

*Data replication sets are available in Harvard Dataverse at: https://doi.org/10.7910/DVN/B6QSZ1

Charli Carpenter (charli.carpenter@gmail.com) is a professor in the Department of Political Science and Legal Studies at University of Massachusetts-Amherst, specializing in international law and human security. She is the author of three books and numerous articles on human rights and the laws of war and has consulted for the UN, State Department, Department of Defense, and human rights NGOs. Her writings appear in the New York Times, Washington Post, Foreign Affairs, Foreign Policy, *and* The American Prospect. *In addition to teaching and research, Professor Carpenter spends time parenting, arguing with concentration camp guards, whitewater kayaking, snowboarding, and rescuing spiders. She tweets @charlicarpenter.*

Alexander H. Montgomery (ahm@reed.edu) is an associate professor of political science at Reed College. He has published articles on theorizing and dismantling proliferation networks; the perils of predicting proliferation, intelligence, and the future of security studies; and the effects of social networks of international organizations on power, interstate conflict, and economic sanctions. He is also coeditor of the Oxford Handbook of Political Networks. *His research interests include political organizations, social networks, weapons of mass disruption and destruction, social studies of technology, and interstate social relations.*

Alexandria J. Nylen (anylen@umass.edu) is a PhD candidate in the Department of Political Science and Legal Studies at the University of Massachusetts-Amherst. Her research interests include international law, critical security studies, and transnational advocacy networks. She has written articles examining the legal and normative implications of remote warfare, such as the use of armed drones in counterterrorism operations. Nylen is currently writing a dissertation on the relationship between issue framing and transnational network connectivity, using advocacy against armed drones as a case study.

We are grateful to the participants in University of Massachusetts-Amherst's Approaches and Methods workshop in February 2020 and at the ISA International Security Studies Section Conference panel on Nuclear Norms in October 2019, to a number of anonymous reviewers, and to generous banks of written comments by Jeff Kaplow and Kevin Young. All flaws and errors are our own.

doi:10.1017/S1537592720002376

© The Author(s), 2020. Published by Cambridge University Press on behalf of the American Political Science Association

https://doi.org/10.1017/S1537592720002376 Published online by Cambridge University Press

humanitarian law and then ask them to express a preference whether or not to do so. From these sorts of studies, some scholars have begun to infer that international laws and norms are weaker than thought, with wider media taking notice.[1]

As an example, in 2017, prominent scholars at Stanford University and Dartmouth College published a now-famous survey experiment to test Americans' sensitivity to the laws of war (Sagan and Valentino 2017). They asked Americans whether, under a fictional scenario involving a war with Iran, they would prefer to conduct an air strike that would kill 100,000 or more Iranian civilians to end the war or to continue the ground war with 20,000 more American military deaths. A slight majority of Americans preferred the strike under most conditions, a result that was highly disseminated and publicized. It was the most frequently downloaded article from *International Security* that year, but it was not only read by scholars. Stanford University put out a press release, and the researchers were interviewed on numerous media spots about their finding. The result was publicized in *Bloomberg, Slate,* and *Foreign Policy*, among many other venues.[2] Most of the reporting focused narrowly on the belief that Americans are supportive of the first use of nuclear weapons, rather than on the conditions that drove public willingness to kill 100,000 civilians. Although a subsequent study by Carpenter and Montgomery (2020) showed that this finding depended not on Americans' insensitivity to international norms and laws but on framing effects embedded in the experiment, many citizens and stakeholders who had read this reporting would likely believe that their fellow Americans favored terrible war crimes and that the nuclear taboo and civilian immunity norm were weak.

What effect do such beliefs have on the actual strength of codified international legal norms, as understood by the public or by political elites? Indeed, what is the relationship between efforts to measure the effects of international legal norms through survey experiments and the resilience of those norms themselves? New research suggests that survey questions, depending on how they are worded and framed, may not simply measure the impact of international law and norms on public opinion but themselves may also affect public sensitivity to these laws—or even citizens' understanding of the content of the laws themselves (Nylen and Carpenter 2019). If so, this has important implications for well-intentioned scholars using survey methods, whether experimental or not, to analyze public opinion on the laws of war.

Using augmented replications of Sagan and Valentino's (2017) survey experiment, we examine whether *merely being invited to express a preference for war crimes* in survey settings has a negative impact on Americans' understanding of US legal and ethical obligations in war.[3] We also examine citizen preferences on the same Iran scenario question to see whether they shift in meaningful ways when they are given the results of the previous survey experiment first.

Our findings are troubling: Americans who answered the Iran Scenario question, were more likely than others to subsequently express a belief that directly and intentionally attacking the civilian population is sometimes permissible under the laws of war, even if they opposed the strike. This finding is important because it is belief in the constraints of international law that offset Americans' willingness to target civilians (Wallace 2019). Respondents are also more likely to prefer to engage in war crimes if told that most other Americans would do so, suggesting that the circulation of such survey results by the media could amplify the erosion of prohibition norms. In short, we find survey experiments such as these, at least in the US context, undermine respondents' understanding of legal norms in the moment they are being measured, and the reported findings can affect the constraining power of those same legal norms on the wider public conscience.

We first discuss the trend toward experimental survey methods in studies of international law and outline the basis for our hypotheses. In the following section, we outline our research design and findings, using both qualitative and quantitative data from a survey experiment on saturation bombing of Iranian civilians. We conclude with suggestions for future experimental survey design in international relations—suggestions that we ourselves intend to take up in any future experiments of this type.

The Martens Trap: Measuring versus Constructing the "Public Conscience"

Research has shown that political elites are at least somewhat sensitive to public attitudes on the use of force and that public attitudes are at least somewhat shaped by international legal norms (Eichenberg 2005; Holsti 2009). It is thus believed that international law works its effects in part through public opinion (Simmons 2009). Indeed, the importance of public opinion makes an appearance in international humanitarian law itself: Russian diplomat Frederich Martens inserted a particular clause into the 1899 Hague Convention that refers states to the "dictates of the public conscience" as a guide where codified law is incomplete (Meron 2000). Norm advocates have invoked this "Martens Clause" as a rationale for measuring public opinion on war law in the service of arguments against certain state behaviors in international affairs (Human Rights Watch 2012).

Accordingly, scholarship on the laws of war has taken a turn toward the experimental, aiming to capture the "public conscience" and exploring how to explain variation in Americans' sensitivity to international laws and norms regarding the use of force (Chilton 2015; Press, Sagan, and Valentino 2013; Wallace 2013). Understanding these dynamics is important, and experimental surveys provide, as Adam Chilton and Dustin Tingley (2013)

https://doi.org/10.1017/S1537592720002376 Published online by Cambridge University Press

argue, ways to correct for important drawbacks in other methods of studying law's effects. The literature is expanding and has led to scholarly debates. Some scholars argue that Americans are fickle and easily persuaded to disregard international legal norms (Conrad et al. 2015; Horowitz 2016). Others find that these norms are robust and effective or that Americans are likelier to heed human rights organizations than their own government when it comes to violating the law (Chilton 2015; Kreps and Wallace 2016; Walsh 2015). Regardless of the causal argument, these studies assume that one can measure the strength of war law using survey experiments without affecting people's understandings of that law.

Yet scholarship on public opinion polls emphasizes how the use of surveys can influence the results and, when disseminated, influence public opinion as well (Druckman 2001). For example, Sarah Kreps (2014) suggests that framing effects in polls on drones result in erroneous data because of the assumptions embedded in the questions and that the media top-lines from those polls become part of the legitimating narrative for targeted killings. Her research shows that pollsters can inadvertently manufacture and inflate support for drone strikes through question framings that discount international law (see also Ron, Lavine, and Golden 2019). Following Kreps, we suspect that the characterization of international law in a wide variety of polls constitutes an important site for citizen socialization into understandings of its content and applicability, particularly in cases where the law is unfamiliar or contested in the public debate.

Indeed, the act of expressing an opinion in a survey may be a *more* powerful socializing device than reading a newspaper, viewing a public service ad, hearing a political speech, or reading an NGO report (Barabas and Jerit 2010, 238). Gaines, Kuklinski, and Quirk (2007, 6) argue that although survey respondents' stated preferences often do not last long beyond the experiment itself, the process of participating in a survey that leaves them with erroneous impressions about facts can cause "permanent changes in their levels of knowledge" (see also Luskin, Fishkin, and Jowell 2002).

When it comes to public knowledge about international legal norms, these changes could potentially be detrimental: Nylen and Carpenter's (2019, 143) analysis of 200 public opinion polls on torture and targeted killings finds that they are often "structured discursively in ways likely to discount or weaken, rather than strengthen, respondents' understanding of the essential character of important international laws and legal debates."

If this is true of public opinion polls, could this critique also apply to survey experiments in political science? Consider a typical survey experiment aiming to explore the effects of international law. In that experiment, a control group reads a fictional scenario on an international issue, and a treatment group reads the same fictional scenario—only with a sentence added that specifies which course of action would or would not be consistent with international law. This type of scenario already constitutes a form of pedagogical hypothetical similar to a war game. The respondent is asked to imagine how reality works given a scenario and, from there, to signal a policy preference. Daniel and Musgrave (2017) demonstrate that imagined experiences, produced by exposure to immersive fiction, stimulate the same cognitive processes that experiences in the real world provoke—and can replace accurate information in the minds of those exposed.[4]

Indeed, the structure of the experiment may socialize the various respondents into an understanding of the very international law whose effects are being measured. This could occur in a part of a survey frame that Lipari (2000, 201) refers to as the "preface" to a survey question: a statement that often contains facts or relevant parameters and precedes the question. The preface "may not only focus and set the stage for questions, but also contain misinformation that perpetuates ideologically charged distortions or misconceptions." Through using a preface, the researcher is thus not simply priming the respondent to test an effect but is also explicitly *educating* the respondent as to what the law is and what it says. Nylen and Carpenter (2019) suggest three significant ways in which exposure to the treatment may either dispel or create substantive perceptions, or misperceptions, about the nature of the law.

First, the nature of the treatment question signals *which* international laws may matter. In cases where multiple international instruments could be relevant or where there are multiple avenues of legal debate, framing the debate in one way versus another can be politically significant. Sarah Kreps (2014), for example, critiques pollsters for incorporating contested government claims, such as the combatant status of terrorists; yet her own survey experiment also incorporates the contested claim that war law (where combatant status matters) is the appropriate legal frame for drone strikes that often happen outside hot battlefields. Similarly, James Walsh's (2015) study of US civilian casualty aversion used scenarios set in Pakistan and Yemen, whereas no state of armed conflict exists between the United States and either country. Although his experiment is sufficient to test his theories about American casualty aversion, it also incorporates contested assumptions about which set of international laws or policy tools is appropriate.

Second, legal treatments are an opportunity to explain to the respondent what international law says: that is, what is prescribed or proscribed by the law.[5] Sometimes, for the treatment group at least, this communicates accurate and relevant information. For example, in the preface to his survey questions on torture, Wallace (2013) defines torture as "severe pain or suffering," which is consistent with the treaty definition. To a public poorly versed in international law, participating in this survey may provide a

https://doi.org/10.1017/S1537592720002376 Published online by Cambridge University Press

significant moment of knowledge dissemination, one that may resonate with the participant long after it is completed. In other cases, however, survey experiments provide a partial or even erroneous understanding of international legal norms. In Valentino and Weinberg's (2017) study of whether genocide language affects respondents' willingness to intervene to stop mass killing, participants are exposed to "experts" characterizing such killings as either "genocide" or "intentional killing of civilians" and then asked not only their policy preferences but also whether they think the killings *deserve* the label genocide. However, they are not provided any information on the legal definition of genocide. Although one could argue that Valentino and Weinberg meant to capture colloquial understandings rather than legal understandings, we wonder about the side effects on respondent understandings of important legal concepts and how they may shape normative debate.

Third, survey experiments often prime respondents to describe under what conditions they would support violating an international norm/law. This can affect understandings of the substance of the norm, because as Shannon (2000) explains, the content of international legal norms includes not simply prescriptions/proscriptions but also *parameters*: conditions under which they are meant to apply or not. International laws vary by how many "exceptions to the rule" they may have: for example, the resort to force is legal under some conditions, but torture and rape are legal under none. By prompting respondents, with respect to an absolute prohibition norm, to state the conditions under which they would support breaking that norm, researchers deliberately invite them to misunderstand the norm itself. The essential intersubjective character of norms (Jurkovich 2019) means that manipulating international law as a variable may give uninformed populations the impression that prohibitions are more flexible than they are.

Meanwhile, in most control groups, international legal considerations are absent. In two separate studies on whether international law influences attitudes toward torture, Geoffrey Wallace asks respondents to imagine a hypothetical scenario in which captured combatants might be tortured in order to extract potentially helpful information. Wallace improves on public opinion polls by explaining to his treatment group that such methods would violate international law and explains that the methods would constitute torture under the relevant treaties (2013, 2014). In neither of these studies, however, is the control group exposed to information on the status of these acts under international law, even after the experiment ended. A study by Adam Chilton and Mila Versteeg (2016) exhibits a similar difficulty: whereas the treatment group is exposed to some information on the status of torture in international law, the control group is not. Depending on their fluency in international law at the outset of the experiment, such participants leave the experiment uninformed and perhaps more poorly informed than before.

Rather than being a "null," Carpenter and Montgomery (2020) have argued this is a kind of priming in itself that cues the participant to believe the law is irrelevant or nonexistent: what they call *priming by omission.* Indeed, entire experiments using scenarios in which international law is highly relevant have been undertaken without reference to international legal obligations in either the control *or* treatment groups (Gartner 2008; Horowitz 2016; Tomz and Weeks 2013; Valentino and Weinberg 2017). For example, in a study on torture, Conrad and coauthors (2015) offered respondents "ticking time bomb" scenarios with varying tactical primes but did not include references to the categorical illegality of torture. Similarly, Horowitz (2016) conducted a study on attitudes toward autonomous weapons while varying the effectiveness of the weapons, their operational necessity, and the number of US troops saved, but did not ask respondents to weigh these factors against the legality or wisdom of outsourcing targeting decisions to machines or against the possibility of civilian deaths. We are aware of two other survey experiments on nuclear strikes that do not mention the illegality of indiscriminate weapons, each focusing instead on practical, strategic, and ethical concerns (Press, Sagan, and Valentino 2013; Rathbun and Stein 2019). Through omission, participants in these studies are "primed" to discount international law, which may not only affect the research findings but also influence their understanding of the relevant factors guiding analogous policy decisions in the real world.

Although it may be argued that survey experiments have little overall effect on wider public understandings because they reach so few respondents, we also hypothesize these effects may be amplified when they are publicized through the media and among beltway stakeholders. As Kreps (2014, 3) pointed out in her analysis of drone polling, "the public's enthusiasm has become part of the political narrative and shaped the boundaries of the conversation." Nylen and Carpenter (2019, 160) hypothesize that the circulation of survey top-lines in mass media "greatly expands their audiences and penetrates the US public discourse, possibly constituting a key linking mechanism between individual polls and wider political understandings." We suspect the same may be true of widely publicized survey experimental research by political scientists.

For example, we analyzed media reactions to Sagan and Valentino's (2017) article and identified 30 discrete media responses. Several were reprinted in multiple news sources, with a slightly disproportionate amount of coverage coming from what MediaBiasChart codes as the "hyperpartisan right-wing" media. For example, *Daily Caller*, a source that falls into MediaBiasChart's "Red Rectangle,"[6] published a piece with the title "Americans Wouldn't

https://doi.org/10.1017/S1537592720002376 Published online by Cambridge University Press

Really Mind if We Just Nuked Somebody," describing the findings as if they were a good thing[7]:

> Americans are much more comfortable with the idea of solving global problems with nuclear weapons, according to a new study. The U.S. faces threats from Russia, China, North Korea, and Iran, as three of these countries have developed weapons systems capable of delivering nuclear payloads to U.S. cities, and one is a sponsor of global terrorist activities and a rising nuclear catastrophe waiting to happen. With all options on the table, the Trump administration is reportedly considering the possibility of a military option in North Korea, with some considering similar options in Iran.

According to Nylen and Carpenter's research on media-based public opinion polls, both the survey effects and the media effects may have a cumulative impact on wider understandings of international norms. Political communication research demonstrates that individuals may not seek out conversations with their peers on controversial political issues for fear of confrontation or uncomfortable social situations (Mutz 2002). This creates a situation in which individual poll respondents are insulated from outside competing elite and non-elite opinions, making it easier to respond to subtle framing effects with counter-humanitarian preferences. The distribution of counter-humanitarian top-lines could create an impression among stakeholders or the public that norms are weak, further socializing individuals into considering extreme measures through social cues and the lens of dramatic threat misperception (Kertzer and Zeitzoff 2017).

If borne out, these hypotheses would pose a serious challenge to the burgeoning literature that uses survey experiments to test international law effects. If nothing else, researchers such as ourselves have a methodological interest in minimizing or controlling effects of the research on our objects of study. In our view, we also have an ethical obligation to ensure that any adverse effects on public understandings, particularly those known to enable mass violence, be noted and minimized. In the next section, we discuss how we tested the impact of survey experiments on international law understandings using a survey experiment to explore these potential side effects empirically. We find support for these hypotheses generated in the critical political communication literature and argue this is a cautionary tale for IR scholars.

Research Design and Findings

To test the impact of survey experiment participation on respondent understandings of international law, we reanalyzed a dataset created as part of an augmented replication of Sagan and Valentino's famous (2017) survey experiment. Sagan and Valentino had surveyed 780 Americans about whether they would agree to an airstrike on a city in Iran that would kill 100,000 or more civilians, if it meant they could end a war and save the lives of 20,000 US troops. Sagan and Valentino's respondents read a fictional news story about a war in Iran that began with a surprise Iranian attack on a US aircraft carrier, leading to a costly ground war. Respondents were told the president was considering a proposal from the Joint Chiefs to conduct a "shock" strike on an Iranian civilian city with the idea of forcing an Iranian surrender. There was no mention in the vignette that this would violate the Geneva Conventions (as might be indicated by a line in the fictional news story reporting opposition to the idea from human rights NGOs). Instead, respondents were cued to believe that the US military thought intentionally killing 100,000 civilians was a permissible option: the vignette stated, "The Chairman of the Joint Chief of Staff delivered a report to the President including an option called the 'shock strategy.'" Respondents were then asked to choose how strongly they would lean toward one of two options. The question read:

> "Given the facts described in the article, if you had to choose between launching the strike against the Iranian city or continuing the ground war against Iran, which option would you prefer?"
>
> STRONGLY PREFER TO CONTINUE GROUND WAR — STRONGLY PREFER TO LAUNCH STRIKE
>
> 1 2 3 4 5 6

In a paper published earlier this year Carpenter and Montgomery (2020) replicated this experiment on an additional 2,500 Americans—with 10 different additional treatment groups to better explore both the effects of international law and the potential framing effects in the original experiment (also see Carpenter and Montgomery 2019). In that paper, we found that respondents were actually quite sensitive to international law knowledge, information, and priming in ways that increased their willingness to risk US soldiers to save civilians. We concluded that the earlier support for a strike had been inflated, because respondents had been primed to discount ethical reasoning by a fictional scenario that expressed a dilemma in purely tactical terms.

In Carpenter and Montgomery (2020), we were able to examine the effect of international law on respondent answers because, unlike Sagan and Valentino's original experiment, respondents were not only asked about their strike preference in the survey. Respondents also received a bank of questions independently about whether they believed it was *legal* to target civilians directly in a war. For example, respondents were asked this question:

> As far as you know, which statement is more correct?
>
> ➢ Under international law, it is sometimes permissible to directly and intentionally attack the civilian population in war.
>
> ➢ Under international law, it is never permissible to intentionally and directly attack the civilian population in war.

Some of the respondents received this question before being asked whether to bomb Iran, and some received it

https://doi.org/10.1017/S1537592720002376 Published online by Cambridge University Press

after: this study design not only enabled us to isolate how much understandings of the content of international law affected expressed preferences but also made it possible to analyze the reverse: how much expressing a preference for or against a war crime affects citizen understanding of *what is legal.* Although we did not explore this latter question in that earlier paper, a reanalysis of the same data allows us to do so here. Thus we test the first of Nylen and Carpenter's (2019) hypotheses:

> (H1): *Being invited to express an opinion about whether to commit a war crime may affect respondents' understanding of the relevant international law itself.*

Does Being Invited to Endorse War Crimes Affect Understandings of What Is Criminal?

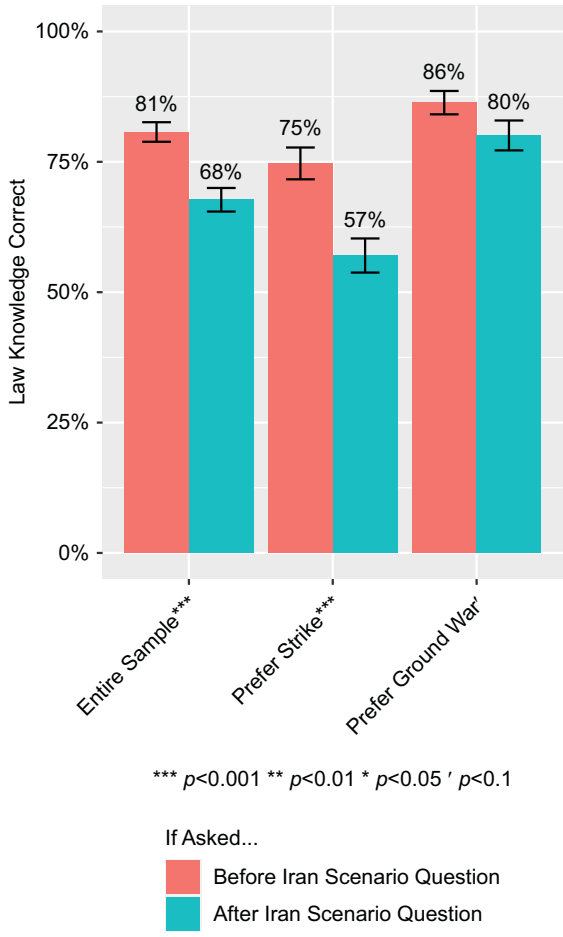
Our findings on H1 and its implications are striking and confirm Nylen and Carpenter's (2019) Hypothesis 1: the percentage of people giving the correct answer to the "Law Knowledge" question dropped dramatically—from 81% to 68%—when the Iran Scenario question was asked first (see Figure 1).

Our initial intuition was that this drop was caused by motivated or consistency bias. Perhaps those who were first asked the survey question and were supportive of saturation bombing would believe (or be motivated to tell themselves) that international law allows for certain exceptions (Kunda 1990). Indeed, this is precisely how norms work: when actors believe in a norm but are motivated to violate it, they often look for ways to argue that the act in question is actually in some way consistent with the norm or that the norm does not apply in this particular situation.[8] This kind of reasoning may indeed have had some influence, because the difference between the "before" and "after" groups was largest for those respondents who signaled they preferred the strike. Those expressing a preference for saturation bombing the city answered the Law Knowledge question correctly 75% of the time when it was asked before the "Iran Scenario" question, but correct responses dropped from 75% to 57% if they were asked about the legality of targeted civilians after already expressing a preference for bombing the city.

More disturbingly, however—and contrary to the motivated reasoning thesis—we observed a similar, if smaller, detrimental effect on correct answers to the "Law Knowledge" question among those who did *not* prefer the strike. Of these, 86% answered the "Law Knowledge" question correctly if it was asked first, compared to only 80% if they were asked about the legality of targeting civilians after being asked to consider whether to strike the city, although this effect was only significant at the $p < 0.1$ level. Thus, even among those inclined to oppose saturation bombing of civilians for ethical or political reasons (in which case motivated reasoning would suggest the opposite finding), simply being asked to consider expressing a preference regarding a war crime seems to reduce some respondents' understanding of whether the act in question is actually prohibited under international law. And this effect is strong enough to counter any reverse-motivated reasoning, where those opposing the strike might be inclined to insist the law was on their side.

Simply put, when respondents are invited to consider an act that runs counter to the laws of war, with no information about *whether* this is legal or ethical but rather only a set of tactical primes suggesting it would be beneficial, they are likelier not only to choose the war crime, as demonstrated in Carpenter and Montgomery's original (2020) paper, but also to leave the experiment believing that such an act *is* legally permissible. The very

Figure 1
Percentage of Respondents with Correct Law Knowledge



Note: Significance testing indicates a difference of means test. Regression with interaction terms shows the same result. We also tested a variety of demographic variables; in each case, we found that the interaction terms were insignificant, indicating that the treatment is not limited to a certain subset of the population.

https://doi.org/10.1017/S1537592720002376 Published online by Cambridge University Press

act of responding to a thought experiment of this type appears to have a socializing effect on respondents, potentially inoculating them against the stopping power of international norms.

Does Media Coverage of Norm "Weakness" Further Weaken Norms?

Building on Nylen and Carpenter's research, we also wanted to know whether knowledge of earlier survey results implying that norms were weak (whether accurately or not) could pose an additional weakening effect on norms. One of the 10 experimental treatment groups embedded in Carpenter and Montgomery's (2020) augmented replication of Sagan and Valentino's (2017) data enabled us to test a second hypothesis:

> (H2): *Knowledge of Sagan and Valentino's earlier survey result, implying that many Americans would choose to target civilians, would weaken the constraining effect of international law on expressed preferences.*

In the treatment group in the Carpenter and Montgomery survey from which we derive our finding, respondents received the same information as every other treatment group with one exception: Sagan and Valentino's original question was rephrased as follows, to incorporate information about Sagan and Valentino's *original* finding.

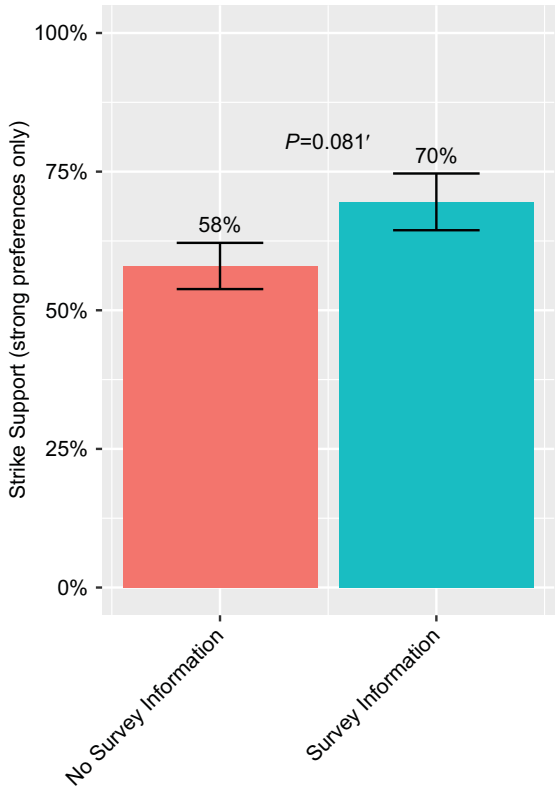
> "Given the facts described in the article, [previous surveys have shown that a majority of Americans would prefer to launch the strike instead of continuing the ground war. If] you had to choose between launching the strike against the Iranian civilians or continuing the ground war against Iran, which option would you prefer?"[9]

STRONGLY PREFER TO CONTINUE GROUND WAR			STRONGLY PREFER TO LAUNCH STRIKE		
1	2	3	4	5	6

As predicted, we found that providing information from the previous survey result that favored bombing the city itself *further* eroded restraint against targeting civilians (see Figure 2).

We arrived at this conclusion by comparing the *ratio* of "strike" to "don't strike" answers for *respondents who felt strongly* one way or another—demonstrated by choosing either a "1" or a "6" on the scale—and dropping out those respondents who registered weak preferences. We chose this approach because respondents who chose more ambivalent answers were not registering preferences so much as protesting against the absence of a "Don't Know" button.[10] Of all the individuals who got no survey information and who felt strongly about striking (6) or not striking (1), 58% chose the strike. This number—the ratio of those who choose to strike the city out of all respondents who feel strongly one way or another—jumps to 70% in the group that *had* information about the previous survey

Figure 2
Effect of Providing Information that "a Majority of Americans Preferred a Strike in a Previous Survey" on Strong Strike Preferences (n = 259)



finding. This increase is statistically significant ($p = 0.081$). Overall, strong support for the strike (6) increased from 21% to 26%; strong support against the strike (for the ground war) fell from 15% to 11%.

We found further evidence for the impact of this effect by examining respondents' explanations for their answers, comparing the open-ended comments in the treatment group who heard about the earlier survey to the comments of those who did *not* hear about the earlier survey before giving their answer. For those respondents who leaned toward the strike, we coded both sets of comments qualitatively to understand how participants justified their answers.[11] We found a range of explanations for striking the city, including that anything was justifiable to win a war; or that the situation seemed like the Hiroshima scenario and therefore must be justified; or that American lives must come first; or that the strike would actually protect both civilians and combatants by ending the war sooner.

Comparing those respondents who received the information about the earlier survey to those who did not, we found insignificant differences on most of the codes, but we did find a significant difference in two codes:

https://doi.org/10.1017/S1537592720002376 Published online by Cambridge University Press

"Americans First" and "Protect Civilians." "Americans First" is a code that aims to capture nationalist sentiment prioritizing US lives or interests above the protection of foreign civilians. For example, one respondent explained his answer this way: "I would rather see other people killed than more of my fellow Americans that are protecting us." Even though some participants acknowledged that civilian lives should be protected as well, when forced to choose between two terrible options as Sagan and Valentino constructed the original survey, many wrote comments like this: "Airstrikes are horrific… the collateral damage is irresponsible. But, all things considered, airstrikes lessen US casualties."

Another interesting, less frequent pattern was a *positive* attitude toward protecting civilians, even among those who leaned toward the strike. Some respondents whose remarks were coded as "Protect Civilians" believed that, if the strike ended the war, then it would ultimately save civilian lives: "They might surrender faster, resulting in fewer military deaths and fewer civilian deaths." Others indicated a faith in the US military to evacuate the city before the bombing: "I know most citizens of that country are told before air strikes. So those people can move away from the area," one typed. Some respondents simply wrote that they felt torn between competing ethical values: "I would rather keep our soldiers safe, but hate the idea of killing innocent women and children."

We found the number of respondents who exhibited an "Americans First" mentality in their answers jumped from 38% to 45% among the treatment group who heard about the earlier survey finding. In other words, hearing that a majority of ones' fellow Americans would prefer the strike, even without knowing what their *rationales* for that strike would be, seemed to exacerbate a desire to prioritize American lives over Iranian lives, even among those who knew it meant violating the laws of war. Moreover, among this treatment group, the moderating effect of concern for protecting civilians dropped in half, from nearly 8% of lean pro-strike comments in the control group to 4% of lean pro-strike comments in the treatment group.

Discussion and Conclusion

Our results show that being asked to express a preference regarding a war crime such as targeting civilians, without being given any information about the laws of war, not only predisposes citizens toward behavior that would violate international humanitarian law but also can undermine their understanding of what is permissible under the laws of war. Our findings also suggest that circulating such results through the media can compound the counter-socializing effect of the original experiment. Citizens are sensitive to hearing what their fellow citizens think of the use of force and seem more likely to approve of committing a war crime if they have heard that a majority would also approve of doing so.

These findings are politically important given that new research on norm contestation and decay suggests that norms are at risk when public perceptions of their *concordance*—that is, intersubjective agreement that they are valid—unravel (Sandholtz 2019). Moreover, this can happen over time when actors persistently chip away at the content of a norm by conceptually expanding the scope conditions beyond which it is understood not to apply (Deitelhoff and Zimmermann 2019, 11). Opinion polls on public (mis)understanding of norms may construct and then carry those meanings to the wider public, having a recursive effect on norm decay.

This leaves analysts of US public opinion and the laws of war, such as ourselves, in an ethical and methodological quandary. On the one hand, citizen morality on the use of force matters politically, so it is important to understand it. On the other hand, researchers invested in valid findings must aim to avoid affecting our subject matter in either direction as it is being measured. Further, we would argue that researchers like ourselves also have a moral obligation to avoid influencing public opinion in ways that are likelier to remove constraints against war crimes—because we know that it is partly citizens' belief in the illegality of certain acts that reduces support for them (Wallace 2019). This perhaps applies most of all to researchers such as ourselves who study mass violence, much as war correspondents are argued to have particular ethical obligations to report from conflict zones in ways that disseminate facts while avoiding the pitfalls by which the media can be used to inflame conflict situations (Galtung and Ruge 1965; Lynch 2015; Lynch and McGoldrick 2005).

Accordingly, we outline three ways that we, as experimental researchers interested in international law, might leverage the insights of this study to create improved research designs *and* facilitate the study of international law effects through experimental methods. We invite other researchers both to join us and to assist us in brainstorming additional practical solutions.

First, we think an area worthy of deliberation among researchers is how to offset any confusion by respondents on the status of international law documented here. For our part, in future studies that we conduct on attitudes toward international law, we pledge to provide all respondents, even those in the control group, the correct information on the status of international law—*after* they answer the question of interest to the dependent variable. Since earlier research shows that accurate knowledge of war law reduces support for war crimes, this approach would potentially counteract any conditioning effect of the experiment (Carpenter and Montgomery 2020; Wallace 2019).[12] In short, given the potential for such experiments to misinform and mislead, we will aim to follow standard IRB practice and disciplinary ethical guidelines for experiments involving deception—even

https://doi.org/10.1017/S1537592720002376 Published online by Cambridge University Press

when deception is inadvertent rather than intentional (Yanow and Schwartz-Shea 2016). We hope other researchers will join us in doing so.

Second, although we acknowledge the value of survey experiments for understanding some dynamics in moral decision making, we also encourage scholars of international law to think more broadly about how these experiments can be designed both to offset ethical quandaries and garner more sophisticated methodological findings. We have found that the use of open-ended survey questions allows a more nuanced analysis than is possible through highly delimited closed-ended answers to experimental survey scenarios (Schaffer 2014). Qualitatively analyzed open-ended questions can allow participants to elaborate and even change their minds as they develop their answers, which, as Hochschild (1981, 238), argues is a more accurate way to capture moral thinking on complex social issues: "People do not make simple statements, they shade, modulate, deny, retract.… These manifestations of uncertainty are just as meaningful and interesting as the definitive statements of a belief system." In Zaller and Feldman's (1992, 585) parlance, therefore, researchers might look not for definitive attitudes but rather *considerations*: "reasons for favoring one side of an issue rather than another." Such an open-ended strategy, in which a respondent is informed of the debate and asked to comment, risks neither priming respondents nor creating any biased or erroneous perceptions.

Third, we think attention must be paid by political scientists to survey vignettes as forms of political communication. Based on our findings here, in our work we will aim to ensure that fictional vignettes accurately reflect the array of political arguments that might be present in a real-world debate, rather than favoring arguments likely to prime participants, such as implying that the Joint Chiefs believe the plan to be permissible.[13] We argue that any vignettes in which citizens are asked to contemplate war crimes include language making clear that these acts would be seen as war crimes by relevant experts, international organizations and NGOs, and some political elites. It is possible to ask respondents how they would feel if the United States were to break international law, without leading them to believe the law itself is permissive (or "too strict") when it comes to important human rights norms. We hope others will join us in this path.

Finally, we should discuss as a research community whether measuring citizen support for war crimes is the best or only way to measure norm strength. The literature on norm contestation and decay suggests a number of ways to evaluate norm robustness, in addition to looking strictly at whether citizens would approve of norm violations. For example, one indicator of norm robustness is third-party reactions to noncompliance (Deitelhoff and Zimmermann 2019; Wiener 2018). Researchers interested in the strength of the nuclear taboo could, for example, ask citizens how their country should respond to violations of that taboo by another actor—rather than asking them to what extent they would approve of the United States violating the taboo itself. Such questions affirm the taboo by inviting participants to consider how to respond, while allowing them freedom to contest the taboo itself unprompted in open-ended questions. Although we are not yet certain whether we would be willing to limit ourselves to this approach, we would welcome a conversation about the relative merits of this and other ideas.

To address the implications of our second finding on media amplification, we suggest that scholars be deliberative in the way in which we publicize experimental survey findings that risk weakening norms, even when the findings are accurate. Given that publicizing a study demonstrating norm weakness could further weaken a norm that holds violence at bay, there may be an ethical duty to frame the finding in ways that are likeliest to minimize harm. For example, a follow-up study by Haworth, Sagan, and Valentino (2019) on nuclear strikes against North Korea showed that two-thirds of Americans would oppose them. The *Bulletin of the Atomic Scientists*' press release emphasized the one-third of Americans who *would* agree with the strikes, and the study argued that this confirmed the immunity norm was weak: "As we have previously found, the U.S. public exhibits only limited aversion to nuclear weapons use and a shocking willingness to support the killing of enemy civilians" (Haworth, Sagan and Valentino 2019, 182). Unsurprisingly, press coverage ran with headlines such as "One Third of Americans Would Support a Nuclear Strike on North Korea, Researchers Say."[14] What might have happened had the press release presented the exact same finding with the opposite framing, emphasizing the majority view and reporting (without emphasizing) the views of the "hawkish minority?" In our case, we would weigh various framings and their potential for benefit or harm as we promoted our work in the public domain, and we again encourage our colleagues to join us.

These ideas and solutions are not exhaustive; we mean to provoke a debate rather than decisively end one. Our wider point (which has been made by many scholars before us in both critical and "quantum" IR scholarship) is that researchers have choices about how we engage with public understandings of the norms we study, as well as an ethical obligation to engage responsibly—even if we adopt the fiction, in our methodologies, that we are apart from the worlds we study (Cox 1981; Wendt 2015).

This call for scholars to consider the real-world political implications of social science survey research extends a robust body of critical literature on reflexivity and the sociology of knowledge in international relations (Haraway 1988; Vitalis 2015). Broadly defined, being reflexive "is to actively 'turn or bend back,' to take account of the self in relation to other subjects and objects," which offers IR scholars the opportunity to potentially "tease out

https://doi.org/10.1017/S1537592720002376 Published online by Cambridge University Press

the subtle ways in which the field may be informed by larger institutional and discursive structures of power" (Eagleton-Pierce 2011, 806). Over the past three decades, the "reflexive turn" has taken on two primary perspectives: a meta-theoretical, "epistemic-normative" one and, increasingly, an interest in extending reflexive practices to empirical IR (Ackerly and True 2008; Hamati-Ataya 2013).

Our methodological suggestions here—to integrate open-ended questions into survey experiments, to include post-experiment correctives for priming effects, and to reflect on the ways in which researchers influence the intersubjective content of the intersubjective norms they study—offer systematic ways for scholars to integrate reflexive practices into empirical IR and perhaps aid in transcending bifurcations between "'explanatory' and 'constitutive' theory" (Hamati-Ataya 2010, 1080; also see Jackson 2015).

Through these and other strategies, we encourage scholars studying public and elite opinion on the laws of war to heed Robert Keohane's advice on responsible scholarship in his essay, "Political Science as a Vocation": to ask questions and to do so in ways that could conceivably matter in making the world a *better* rather than a worse place—and to be cognizant of and reflective on our power as scientific authorities (Keohane 2009). These are arguments long made by critical and reflectivist IR scholars, ethnographers, and particularly those studying social relations in conflict zones (Ackerly, Stern, and True 2006; Pachirat 2017; Smyth and Robinson 2001). But we note Keohane's remarks because we argue that this sensibility—to reflect on our role in our subject matter and account in ethical ways for the effects of our work on the world we study—should not be limited to critical scholarship. Rather, we see this as a cornerstone of all good *science.*[15] As Kevin Young (2020) puts it:

> The role of moral values aren't absent from this process, but rather are understood to be an integral part of it – with the operative point being that the ultimate goal of scientific knowledge is to produce a set of facts that even someone who did not share our values would find compelling (Jackson 2015: 946)… this capacious understanding of science…suggests that many of the sharp critiques of the "partial perspective" of researchers… and the value-laden nature of concepts and their histories, can be considered as part of the conversation for the improvement of the scientific enterprise, rather than motivations for discarding it."

Indeed, important empirical, even quantitative, work in political science and legal studies has begun to document the partial and even biased perspectives hidden in mainstream scholarship on international law and relations in ways that test (and sometimes confirm) interpretive and critical perspectives, as well as offering mechanisms for scholars to be conscious of and control for these implicit biases (Chilton and Posner 2015; Rathbun 2012).

For conflict researchers using critical, ethnographic or qualitative methods, these practical, personal and political dilemmas are often front and center of the way they orient us as scholars and as people toward the research enterprise: what Sarah E. Parkinson (2018) refers to, in the title of her essay memorializing Lee Ann Fujii, as "being human in political science." Indeed, the late Fujii (2012, 717) encouraged a consideration of ethical quandaries in the conduct, dissemination, and use of all types of political science research: "Producers and consumers of political science research need to contemplate the ambiguous and oftentimes uncomfortable dimensions of research ethics, lest we create a discipline that is 'non-ethical,' or worse, unethical."

Survey experimentalists studying attitudes toward international humanitarian law may be less attuned to these research dilemmas than field researchers in conflict zones, but our findings suggest that the need to reflect on our role in our research and minimize harm is no less imperative, especially when studying (and perhaps subtly shaping) attitudes enabling and constraining war crimes. This imperative—not only to research responsibly but also to *engage* responsibly—becomes even more urgent as we choose when, where, and how to disseminate our findings to stakeholders and the wider public.[16]

At minimum, as survey experimentalists collect data, we must be cautious not to overdetermine the results in ways that undermine the intersubjective protective power of international ethical norms, provide elites a road map to manufacturing consent for war crimes, encourage citizens to whitewash or justify terrible acts, or otherwise lay a foundation for placing vulnerable populations at greater risk of political violence.

Notes

1 For the purposes of this article, we use the term "international law" to refer to the provisions of codified treaty law and international "norms" to refer to the intersubjectively agreed-on ethical standards that are sometimes, but not always, expressed in treaty provisions.

2 See the Altmetric bookmarklet for "Revisiting Hiroshima in Iran," https://www.altmetric.com/details/23376728/news.

3 Although we cannot generalize beyond the US case due to the nature of the sample, it is plausible that this effect may hold in other country contexts as well.

4 Active learning practitioners and analysts who have studied why war gaming works emphasize the way in which hypothetical scenarios engage learning processes (Pauly 2018; Perla and McGrady 2011).

5 Although some (e.g., Hurd 2017) will argue that the "correct" meaning of international law is never settled, this is more true for some norms than for others, in which the content of the legal norm is ambiguous or has many scope conditions. Our research, like much of the work described here, focuses on very concrete

https://doi.org/10.1017/S1537592720002376 Published online by Cambridge University Press

prohibition norms, such as the norm against torture or against targeting civilians. Although it is true that the concept of who is a civilian may be fuzzy, the prohibition against targeting civilians is quite clear.

6 MediaBiasChart codes news sources according to both partisan bias and overall quality, with the partisan bias grid varying (on the horizontal axis) from Hyperpartisan Left to Hyperpartisan Right (through "skews left/right" and "neutral") and with the overall quality grid covering a vertical-axis continuum between "Original Fact Reporting" to "Contains Inaccurate/Fabricated Info." Sources are then categorized to where they fall on a grid, with sources in the "Red Rectangle" containing both extreme partisan skew and propensity toward fake news, or what MediaBiasChart calls "Nonsense Damaging to Public Discourse." "Media Bias Chart: Version 4.0." Ad Fontes Media. Accessed June 23, 2019. https://www.adfontesmedia.com/.

7 Ryan Pickrell, 2017. "Americans Wouldn't Really Mind if We Just Nuked Somebody." *Daily Caller.* August 4. https://dailycaller.com/2017/08/04/americans-wouldnt-really-mind-if-we-just-nuked-somebody/. Accessed June 23, 2019.

8 See Shannon (2000) for an analysis of the political psychology with which actors account for their norm violations by convincing themselves that the act in question falls outside the scope of a norm. See Panke and Petersohn (2012) on how "applicatory justification" can potentially weaken norms by narrowing their scope. See also Hurd (2017) and Búzás (2017) on how law can be used to justify illegal or immoral behavior.

9 Brackets omitted in actual survey.

10 We saw similar, but less statistically significant, results when including more ambivalent answers (2, 3, 4, and 5): including all answers, we saw an increase in preferences to strike the city from 54% to 61% (p = 0.1097, n = 739); excluding only the most ambivalent (those who chose 3 or 4 on a 6-point scale) resulted in an increase in strike preference from 57% to 64% (p = 0.1732, n = 456).

11 Open-ended comments were analyzed through a grounded-theory approach first by a team of two undergraduate coders and then coded independently by these two coders using a code scheme developed by one of the coauthors. The two coders achieved an average inter-rater reliability score of 91.5. Mismatched codes were adjudicated by the researcher for accuracy; ambiguous codes were adjudicated through a consensus coding process. The codebook is available on request.

12 Although it would not seem appropriate for researchers to express moral values in an experiment, the dissemination of factual information seems appropriate because it is a social fact about the world.

13 In reality, as Colin Kahl (2007) details, US military doctrine would make it extremely unlikely that top generals would approve of such a terrible war crime.

14 Simon Denyer, "One Third of Americans Would Support a Nuclear Strike on North Korea, Researchers Say," *Washington Post,* June 25, 2019, p. 4; see also Kelsey Piper, "Americans Are Terrifyingly Supportive of Nuking Civilians," *Vox*, June 26, 2019.

15 By "science" here we adopt Patrick Jackson's (2011) formulation and refer to both explanatory and descriptive work that involves *systematic* claims about social and political arrangements.

16 This imperative, for example, forms the basis for new efforts within the discipline to encourage responsible engagement, such as the Sie Center's Responsible Engagement Board at University of Denver. See Sie Center, "The Three Rs: Rigor, Reliability, and Responsibility," *Political Violence at a Glance*, May 21, 2019. http://politicalviolenceataglance.org/2019/05/21/the-three-rs-rigor-relevance-and-responsibility-in-policy-engaged-research/. Accessed June 21, 2020.

References

Ackerly, Brooke A., Maria Stern, and Jacqui True. 2006. *Feminist Methodologies for International Relations*. Cambridge: Cambridge University Press.

Ackerly, Brooke A., and Jacqui True. 2008. "Reflexivity in Practice: Power and Ethics in Feminist Research on International Relations." *International Studies Review* 10 (4): 693–707.

Barabas, Jason, and Jennifer Jerit. 2010. "Are Survey Experiments Externally Valid?" *American Political Science Review* 104 (2): 226–42.

Búzás, Zoltán I. 2017. "Evading International Law: How Agents Comply with the Letter of the Law but Violate Its Purpose." *European Journal of International Relations* 23 (4): 857–83.

Carpenter, Charli, and Alexander H. Montgomery. 2019. "Americans Want Their Leaders to Obey the Laws of War." *Foreign Policy*, June 27. https://foreignpolicy.com/2019/06/27/americans-want-their-leaders-to-obey-the-laws-of-war./

Carpenter, Charli, and Alexander H. Montgomery. 2020. "The Stopping Power of Norms: Saturation Bombing, Civilian Immunity, and U.S. Attitudes toward the Norms and Laws of War." *International Security* (forthcoming).

Chilton, Adam S. 2015. "The Laws of War and Public Opinion: An Experimental Study." *Journal of Institutional and Theoretical Economics JITE* 171 (1): 181–201.

Chilton, Adam S., and Eric A. Posner. 2015. "An Empirical Study of Political Bias in Legal Scholarship." *Journal of Legal Studies* 44 (2): 277–314.

https://doi.org/10.1017/S1537592720002376 Published online by Cambridge University Press

Chilton, Adam S., and Dustin Tingley. 2013. "Why the Study of International Law Needs Experiments." *Columbia Journal of Transnational Law* 52: 173.
Chilton, Adam S., and Mila Versteeg. 2016. "International Law, Constitutional Law, and Public Support for Torture." *Research & Politics* 3 (1).
Conrad, Courtenay R., Sarah E. Croco, Brad T. Gomez, and Will H. Moore. 2015. "Threat Perception and Public Support of Government Torture." Working Paper. University of California, Merced.
Cox, Robert W. 1981. "Social Forces, States and World Orders: Beyond International Relations Theory." *Millennium: Journal of International Studies* 10 (2): 126–55.
Daniel III, J. Furman, and Paul Musgrave. 2017. "Synthetic Experiences: How Popular Culture Matters for Images of International Relations." *International Studies Quarterly* 61 (3): 503–16.
Deitelhoff, Nicole, and Lisbeth Zimmermann. 2019. "Norms under Challenge: Unpacking the Dynamics of Norm Robustness." *Journal of Global Security Studies* 4 (1): 2–17.
Druckman, James N. 2001. "The Implications of Framing Effects for Citizen Competence." *Political Behavior* 23 (3): 225–56.
Eagleton-Pierce, Matthew. 2011. "Advancing a Reflexive International Relations." *Millennium: Journal of International Studies* 39 (3): 805–23.
Eichenberg, Richard C. 2005. "Victory Has Many Friends: US Public Opinion and the Use of Military Force, 1981–2005." *International Security* 30 (1): 140–177.
Fujii, Lee Ann. 2012. "Research Ethics 101: Dilemmas and Responsibilities." *PS: Political Science & Politics* 45 (4): 717–23.
Gaines, Brian J., James H. Kuklinski, and Paul J. Quirk. 2007. "The Logic of the Survey Experiment Reexamined." *Political Analysis* 15 (1): 1–20.
Galtung, Johan, and Mari Holmboe Ruge. 1965. "The Structure of Foreign News: The Presentation of the Congo, Cuba and Cyprus Crises in Four Norwegian Newspapers." *Journal of Peace Research* 2 (1): 64–90.
Gartner, Scott Sigmund. 2008. "The Multiple Effects of Casualties on Public Support for War: An Experimental Approach." *American Political Science Review* 102 (1): 95–106.
Hamati-Ataya, Inanna. 2010. "Knowing and Judging in International Relations Theory: Realism and the Reflexive Challenge." *Review of International Studies* 36 (4): 1079–1101.
Hamati-Ataya, Inanna. 2013. "Reflectivity, Reflexivity, Reflexivism: IR's 'Reflexive Turn' — and Beyond." *European Journal of International Relations* 19 (4): 669–94.
Haraway, Donna. 1988. "Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective." *Feminist Studies* 14 (3): 575–99.
Haworth, Alida R., Scott D. Sagan, and Benjamin A. Valentino. 2019. "What Do Americans Really Think about Conflict with Nuclear North Korea? The Answer Is Both Reassuring and Disturbing." *Bulletin of the Atomic Scientists* 75 (4): 179–186.
Hochschild, Jennifer L. 1981. *What's Fair? American Beliefs about Distributive Justice*. Cambridge, MA: Harvard University Press.
Holsti, Ole Rudolf. 2009. *Public Opinion and American Foreign Policy*.Rev.ed. Ann Arbor: University of Michigan Press.
Horowitz, Michael C. 2016. "Public Opinion and the Politics of the Killer Robots Debate." *Research & Politics* 3 (1): 1–8.
Human Rights Watch. 2012. *Losing Humanity: The Case against Killer Robots*. New York: Human Rights Watch.
Hurd, Ian. 2017. *How to Do Things with International Law*. Princeton: Princeton University Press.
Jackson, Patrick Thaddeus. 2011. *The Conduct of Inquiry in International Relations*. London: Routledge.
Jackson, Patrick Thaddeus. 2015. "Must International Studies Be a Science?" *Millennium: Journal of International Studies* 43 (3): 942–65.
Jurkovich, Michelle. 2019. "What Isn't a Norm? Redefining the Conceptual Boundaries of 'Norms' in the Human Rights Literature." *International Studies Review*, August 20. https://academic.oup.com/isr/advance-article/doi/10.1093/isr/viz040/5550876.
Kahl, Colin H. 2007. "In the Crossfire or the Crosshairs? Norms, Civilian Casualties, and US Conduct in Iraq." *International Security* 32 (1): 7–46.
Keohane, Robert O. 2009. "Political Science as a Vocation." *PS: Political Science and Politics* 42 (2): 359–63.
Kertzer, Joshua D., and Thomas Zeitzoff. 2017. "A Bottom-Up Theory of Public Opinion about Foreign Policy." *American Journal of Political Science* 61 (3): 543–58.
Kreps, Sarah. 2014. "Flying under the Radar: A Study of Public Attitudes towards Unmanned Aerial Vehicles." *Research & Politics* 1 (1): 1–7.
Kreps, Sarah, and Geoffrey P. R. Wallace. 2016. "International Law, Military Effectiveness, and Public Support for Drone Strikes." *Journal of Peace Research* 53 (6): 830–844.
Kunda, Ziva. 1990. "The Case for Motivated Reasoning." *Psychological Bulletin* 108 (3): 480–98.
Lipari, Lisbeth. 2000. "Toward a Discourse Approach to Polling." *Discourse Studies* 2 (2): 187–215.
Luskin, Robert C., James S. Fishkin, and Roger Jowell. 2002. "Considered Opinions: Deliberative Polling in Britain." *British Journal of Political Science* 32 (3): 455–87.

https://doi.org/10.1017/S1537592720002376 Published online by Cambridge University Press

Lynch, Jake. 2015. "Peace Journalism: Theoretical and Methodological Developments." *Global Media and Communication* 11 (3): 193–99.
Lynch, Jake, and Annabel McGoldrick. 2005. *Peace Journalism*. Stroud, UK: Hawthorn Press.
Meron, Theodor. 2000. "The Martens Clause, Principles of Humanity, and Dictates of Public Conscience." *American Journal of International Law* 94 (1): 78–89.
Mutz, Diana C. 2002. "The Consequences of Cross-Cutting Networks for Political Participation." *American Journal of Political Science* 46 (4): 838–55.
Nylen, Alexandria, and Charli Carpenter. 2019. "Questions of Life and Death." *European Journal of International Security* 4 (2): 142–62.
Pachirat, Timothy. 2017. *Among Wolves: Ethnography and the Immersive Study of Power*. London: Routledge.
Panke, Diana, and Ulrich Petersohn. 2012. "Why International Norms Disappear Sometimes." *European Journal of International Relations* 18 (4): 719–42.
Parkinson, Sarah E. 2018. "The Dignity of Complexity: Being Human in Political Science." *Qualitative and Multi-Method Research* 16 (1): 39–41.
Pauly, Reid B. C. 2018. "Would U.S. Leaders Push the Button? Wargames and the Sources of Nuclear Restraint." *International Security* 43 (2): 151–92.
Perla, Peter P, and Ed McGrady. 2011. "Why Wargaming Works." *Naval War College Review* 64 (3).
Press, Daryl G., Scott D. Sagan, and Benjamin A. Valentino. 2013. "Atomic Aversion: Experimental Evidence on Taboos, Traditions, and the Non-Use of Nuclear Weapons." *American Political Science Review* 107 (1): 188–206.
Rathbun, Brian C. 2012. "Politics and Paradigm Preferences: The Implicit Ideology of International Relations Scholars." *International Studies Quarterly* 56 (3): 607–22.
Rathbun, Brian C., and Rachel Stein. 2019. "Greater Goods: Morality and Attitudes toward the Use of Nuclear Weapons." *Journal of Conflict Resolution*, December 27. https://doi.org/10.1177/0022002719879994.
Ron, James, Howard Lavine, and Shannon Golden. 2019. "No, Americans Don't Support Airstrikes That Kill Civilians, Even When They Target Terrorists." *Washington Post*, June 13. https://www.washingtonpost.com/politics/2019/05/06/no-americans-dont-support-airstrikes-that-kill-civilians-even-when-they-target-terrorists/.
Sagan, Scott D., and Benjamin A. Valentino. 2017. "Revisiting Hiroshima in Iran: What Americans Really Think about Using Nuclear Weapons and Killing Noncombatants." *International Security* 42 (1): 41–79.
Sandholtz, Wayne. 2019. "Norm Contestation, Robustness, and Replacement." *Journal of Global Security Studies* 4 (1): 139–46.
Schaffer, Frederic Charles. 2014. "Thin Descriptions: The Limits of Survey Research on the Meaning of Democracy." *Polity* 46 (3): 303–30.
Shannon, Vaughn P. 2000. "Norms Are What States Make of Them: The Political Psychology of Norm Violation." *International Studies Quarterly* 44 (2): 293–316.
Simmons, Beth A. 2009. *Mobilizing for Human Rights: International Law in Domestic Politics*. Cambridge: Cambridge University Press.
Smyth, Marie, and Gillian Robinson. 2001. *Researching Violently Divided Societies: Ethical and Methodological Issues*. New York: United Nations University Press.
Tomz, Michael R., and Jessica L. P. Weeks. 2013. "Public Opinion and the Democratic Peace." *American Political Science Review* 107 (4): 849–65.
Valentino, Benjamin A., and Ethan M. Weinberg. 2017. "More than Words? 'Genocide,' Holocaust Analogies, and Public Opinion in the United States." *Journal of Human Rights* 16 (3): 276–92.
Vitalis, Robert. 2015. *White World Order, Black Power Politics: The Birth of American International Relations*. Ithaca: Cornell University Press.
Wallace, Geoffrey P. R. 2013. "International Law and Public Attitudes toward Torture: An Experimental Study." *International Organization* 67 (1): 105–140.
Wallace, Geoffrey P. R. 2014. "Martial Law? Military Experience, International Law, and Support for Torture." *International Studies Quarterly* 58 (3): 501–14.
Wallace, Geoffrey P. R. 2019. "Condemning or Condoning the Perpetrators? International Humanitarian Law and Attitudes toward Wartime Violence." *Law & Social Inquiry* 44 (1): 192–226.
Walsh, James Igoe. 2015. "Precision Weapons, Civilian Casualties, and Support for the Use of Force." *Political Psychology* 36 (5): 507–523.
Wendt, Alexander. 2015. *Quantum Mind and Social Science: Unifying Physical and Social Ontology*. Cambridge: Cambridge University Press.
Wiener, Antje. 2018. *Contestation and Constitution of Norms in Global International Relations*. Cambridge: Cambridge University Press.
Yanow, Dvora, and Peregrine Schwartz-Shea. 2016. "Encountering Your IRB 2.0: What Political Scientists Need to Know." *PS: Political Science & Politics* 49 (2): 277–86.
Young, Kevin L. 2020. "Progress, Complexity and Science: Moving from Alienated to Engaged Pluralism." Working Paper, University of Massachusetts-Amherst.
Zaller, John, and Stanley Feldman. 1992. "A Simple Theory of the Survey Response: Answering Questions versus Revealing Preferences." *American Journal of Political Science* 36 (3): 579–616.

https://doi.org/10.1017/S1537592720002376 Published online by Cambridge University Press