Antinormative Messaging, Group Cues, and the Nuclear Ban Treaty

Article *in* The Journal of Politics · January 2022

DOI: 10.1086/714924

3 authors, including:

Stephen Herzog
ETH Zurich
40 PUBLICATIONS **393** CITATIONS
SEE PROFILE

Jonathon Baron
Yale University
12 PUBLICATIONS **306** CITATIONS
SEE PROFILE

SHORT ARTICLE

Antinormative Messaging, Group Cues, and the Nuclear Ban Treaty

Stephen Herzog, ETH Zurich
Jonathon Baron, Yale University
Rebecca Davis Gibbons, University of Southern Maine

What types of foreign policy cues are most likely to turn public opinion against a popular emerging norm? Since 2017, the US government has sought to discredit the Treaty on the Prohibition of Nuclear Weapons and its nuclear nonpossession norm among the largely prodisarmament American public. We fielded a national US survey experiment ($N = 1,219$) to evaluate the effects of these elite cues as well as social group cues on public opinion. Our study thus offers one of the first experimental assessments of public attitudes toward nuclear disarmament. We find that both negative government messages and group cues can shift attitudes. Direct exposure to official rhetoric—particularly substantive security and institutional critiques—most effectively increases opposition to the norm. Yet, we observe that all cues have little effect on respondents' existing opposition to nuclear arms. The American population may support eventually eliminating nuclear weapons, but majority backing for immediate disarmament appears far from assured.

In 2017, 122 states adopted a treaty outlawing possession of nuclear weapons, but no state with a nuclear arsenal or protected by extended nuclear deterrence voted in favor. British, French, and US officials immediately responded: "We do not intend to sign, ratify or ever become party to" the Treaty on the Prohibition of Nuclear Weapons (TPNW; Gibbons 2019, 30). Regardless, ban advocates aim to influence governments by turning their domestic audiences against nuclear weapons (Mekata 2018). Literature showing that public opinion can shape foreign policy (Milner and Tingley 2015; Powlick and Katz 1998) suggests that this strategy might yield promising results in the United States. US government efforts to discredit the TPNW also face an uphill battle due to long-standing support for nuclear disarmament among most Americans (Kull et al. 2004; Russett 1990–91).

Can negative messaging sway opinion against the emergent nuclear nonpossession norm? If so, what types of messages are most persuasive to a largely prodisarmament public? Literature on foreign policy cues is mixed, with Guisinger and Saunders (2017) finding that the public follows elite views and Kertzer and Zeitzoff (2017) concluding that group cues often matter more. However, no study disaggregates how negative messaging affects support for international nuclear norms, and most recent nuclear politics experiments explore public support for the use and nonuse of nuclear weapons (see, e.g., Sagan and Valentino 2017). The only experiment to address

Stephen Herzog (stephen.herzog@sipo.gess.ethz.ch) is a senior researcher in the Center for Security Studies, ETH Zurich, Switzerland. Jonathon Baron (jonathon.baron@yale.edu) is an independent researcher and a recent PhD graduate of the Department of Political Science, Yale University, New Haven, CT 06511. Rebecca Davis Gibbons (rebecca.gibbons@maine.edu) is an assistant professor in the Political Science Department, University of Southern Maine, Portland, ME 04103.

This study received an exempt determination per Yale University Institutional Review Board protocol 2000026191. Study preregistration occurred through Evidence in Governance and Politics preanalysis plan 20190806AA. Support for this research was provided by the Whitney and Betty MacMillan Center for International and Area Studies at Yale University, the Japan Foundation Center for Global Partnership, and the Stanton Foundation. Data and supporting materials necessary to reproduce the numerical results in the article are available in the *JOP* Dataverse (https://dataverse.harvard.edu/dataverse/jop). An online appendix with supplementary material is available at https://doi.org/10.1086/714924.

Published online October 21, 2021.

  Published by The University of Chicago Press for the Southern Political Science Association. https://doi.org/10.1086/714924

nuclear disarmament does so as a means to assess the comparative strength of positivist international relations theories (Bell and Quek 2015).

Our study fills this gap on disarmament with a US national survey experiment ($N = 1,219$). We randomly assigned respondents to a single persuasive treatment (or control) from a range of vignettes arguing against the TPNW on security, normative, or institutional grounds, mirroring US government messages. The survey also evaluated a group cue containing a social prime but no substantive objections. The results offer theoretical insights into government strategies to counter emergent norms as well as "trickle-down" group effects.

Both government messages and group cues prompted greater opposition to the TPNW, but priming Americans' distrust of international institutions and security concerns had the strongest effects. While 64.7% of control respondents supported joining the ban, describing it as having weak institutional efficacy yielded a considerable 19.2 percentage point decrease. Discussing its allegedly harmful impact on US national security attenuated support by 17 percentage points. Noting its potential to undermine nonproliferation norms dropped support by 8.6 percentage points. Social group cues reduced support by 8.1 percentage points. On average, government messaging most effectively increased opposition, but in practice, cues often work in tandem. Since group cues have a significant effect on support, official messages may be magnified as people persuaded by rhetoric opposing the ban transmit information through social networks. Yet, all cues had little effect on respondents' existing opposition to nuclear arms. The US public may support eventually eliminating nuclear weapons, but majority backing for immediate disarmament appears far from assured.

EXPERIMENTAL DESIGN

We conducted our survey from August 9 to 12, 2019, using a national sample of the US population balanced on age, gender, and region, recruited by the polling firm Dynata.[1] The experiment, shown in figure B1, proceeded as follows. After consent, subjects provided demographic and political covariates. Next, they read a baseline description of the TPNW, including the number of states that negotiated its adoption and its overall objectives (app. C). Subjects were then randomly assigned to one of four treatment arms—or the control group receiving no further information—before responding to outcome measures.

Treatments

Except for the group cue, interventions modeled US government elite arguments against the TPNW. Table D1 demonstrates that covariates were balanced across the four arms occurring alongside the control, the text of which is in appendix C:

1. *Group Cue*: augmented replication of Kertzer and Zeitzoff (2017) with a figure noting "those who answered other survey questions like you do not support the Ban Treaty";[2]
2. *Security Cue*: statement noting US government opposition due to the goal of eliminating nuclear weapons used for protection against other nuclear powers;
3. *Norms Cue*: statement noting US government opposition due to the potential to subvert norms of the Nuclear Nonproliferation Treaty (NPT);
4. *Institution Cue*: statement noting US government opposition due to the ban's status as a weak international institution lacking enforcement and verification.

External validity

We assessed causal effects within the context of a real-world case. US officials have countered the ban's public visibility with strong condemnatory rhetoric mirroring our treatments. Following its adoption, Ambassador to the United Nations Nikki Haley's press statement warned, "This initiative clearly disregards the realities of the international security environment. Accession to the ban treaty is incompatible with the policy of nuclear deterrence, which has been essential to keeping the peace in Europe and North Asia for over 70 years" (Casey-Maslen 2019, 52). When the International Campaign to Abolish Nuclear Weapons won the 2017 Nobel Peace Prize, the State Department critiqued the ban's institutional efficacy, stating it "will not result in the elimination of a single nuclear weapon" (Keaten and Lewis 2017). As state signatories have risen, officials have invoked norms in—often televised—speeches at think tanks and universities. Assistant Secretary of State Christopher Ford (2018) has alleged that the ban "works at cross-purposes" to the "tried and true institutions of the NPT" that have promoted nuclear security for five decades. Many Americans receive such messages directly or through factual reporting—similar to

1. Our Dynata sample was well balanced on relevant US demographic covariates (app. A; apps. A–K are available online), as was a follow-up sample recruited by YouGov (see Primary Treatment Effects).

2. This treatment used mild deception, so subjects received a debrief found in app. E.

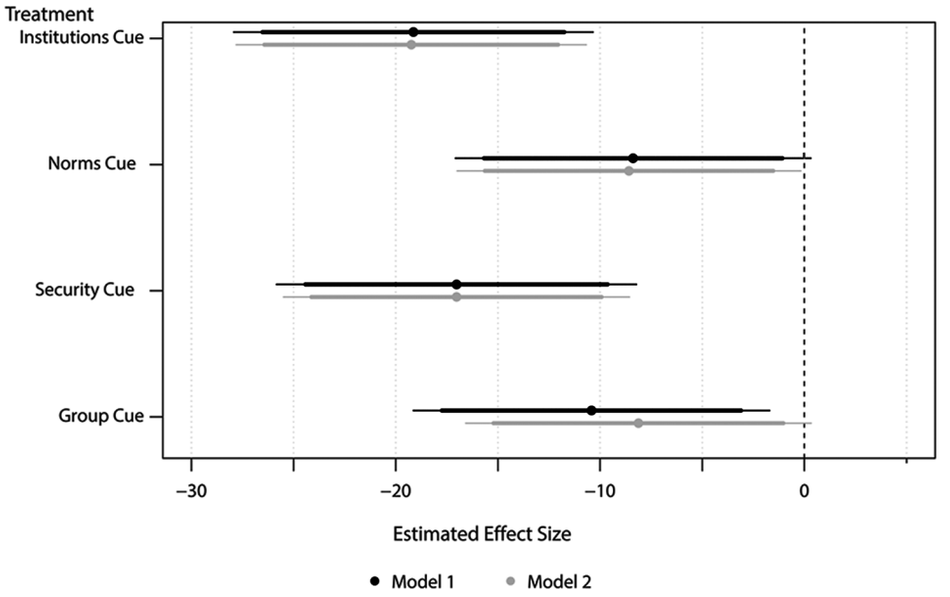

Figure 1. Treatment effects on support for the TPNW. Point estimates of treatment effects measured against the control baseline are presented for model 1 (black) and model 2 (gray), with two-sided 90% confidence intervals (pertaining to one-tailed hypothesis tests) and 95% confidence intervals. Color version available as an online enhancement.

substantive treatment. Others learn about news via friends, family, or partisan sound-bite commentary—akin to group cues and often containing little informational content. Further discussion of our experimental approach for evaluating these messages is provided in appendix C.

Posttreatment outcome measures

The primary outcome measure asked respondents: "Do you think the United States should join the Nuclear Weapon Ban Treaty?" Subjects then completed an attitudinal battery asking their views on three types of statements about nuclear weapons: whether they are an asset or a danger to international peace; whether they are usable as weapons of war; and how subjects view disarmament dynamics like feasibility, verifiability, and implementation speed. Appendix F shows all outcome measures. We wanted to know whether treatment could influence perceptions of not only the ban but nuclear weapons themselves. Deeply penetrating effects might spill over to issues of nuclear arms control, proliferation, deterrence, and force posture. We hypothesized that each treatment would negatively affect support for the ban. Accordingly, we scored subjectively positive responses on a four-point Likert scale with values of -2, -1, 1, and 2; higher numerical coding facilitated directional testing.

ANALYSES AND RESULTS

We conducted our main analyses using ordinary least squares regression models with HC2 robust standard errors (SEs). Model 1 estimated

$$Y_i = \beta_0 + \beta_1 \text{Group Cue}_i + \beta_2 \text{Security Cue}_i + \beta_3 \text{Norms Cue}_i + \beta_4 \text{Institution Cue}_i + \varepsilon_i, \quad (1)$$

or the effect of each treatment relative to the control baseline on each outcome Y (and error ϵ_i)—the difference in means for each treatment group relative to the control. Model 2 estimated

$$Y_i = \gamma_0 + \gamma_1 \text{Group Cue}_i + \gamma_2 \text{Security Cue}_i + \gamma_3 \text{Norms Cue}_i + \gamma_4 \text{Institution Cue}_i + \mathbf{X}_i + \eta_i, \quad (2)$$

where $\mathbf{X}_i$ represents the covariate profile of each respondent i with associated regression coefficients and error term η_i. Results did not differ substantially, and we report on both models in figure 1, although our primary inferential targets are the adjusted estimates in model 2.[3] Because our primary hypotheses posited negative treatment effects, our design was preregistered with the specification that all effects would be tested against the null hypothesis of no effect using one-tailed, lower p-values ($\alpha = 0.05$; see app. G). However, we also present results with two-sided 95% confidence intervals.

Primary treatment effects

In line with our hypotheses, figure 1 displays significant negative effects of each treatment on support for joining the TPNW (see also app. H). As noted above, the baseline level of

3. We report unadjusted estimates in Broader Attitudinal Effects and Subgroup Analysis.

support among control respondents was 64.7%. Support was lowest among respondents in the institution and security cue groups, respectively 45.6% (regression-adjusted estimate: average treatment effect [ATE] = −19.2, $p < .001$) and 47.7% (ATE = −17.0, $p < .001$). These effects were statistically indistinguishable. We estimated that 54.3% of group cue respondents supported the TPNW (ATE = −8.1, $p < .030$), while 56.3% of norms cue respondents supported it (ATE = −8.6, $p = .023$). These effects were not significantly different at the $p < .05$ level, but each was significantly smaller than the institution and security cue effects.[4]

Because Americans may be unfamiliar with the ban, we also performed a follow-up survey to assess priors on the TPNW. This matched representative survey ($N = 2,500$), conducted by the firm YouGov, allowed us to contextualize the importance of the treatment effects. Although only 26.1% of Americans were aware of the TPNW, even unaware respondents revealed strong preferences for disarmament. Around 61% of respondents who had not heard of the ban noted it "sound[ed] like they would support it." Appendix I provides full results. In fact, the proportions of subjects supporting and opposing the ban were similar among both aware and unaware respondents. With this strong evidence of overt and latent support for the TPNW, it is apparent that antinormative messaging can indeed be powerful.

Broader attitudinal effects

Contrasting with the strong effects on the primary outcome of interest, no treatment had a systematic effect on broader attitudes toward nuclear weapons (see table J1). Subjects were, on average, negative to neutral in their views on nuclear weapons. For instance, the average control group response to "Nuclear weapons are dangerous and present a threat to the world" was −1.332 (SE = 0.058), where −2 represented the most negative attitudes. Tellingly, the average control group response to "Some countries will always cheat and disobey nuclear treaties" was also decidedly negative (mean = −1.265; SE = 0.057). Americans already appear to believe that countries are likely to behave deceptively in the context of nuclear diplomacy. Accordingly, we saw moderate endorsement of the idea that "Reducing the number of nuclear weapons over time is safer than immediate nuclear disarmament" (mean = 0.748; SE = 0.076), where 2 represented the most positive attitudes.

Yet, respondents were not indiscriminately negative about nuclear weapons. Baseline attitudes showed modest agreement with "Nuclear weapons help to keep my country safe" (mean = 0.332; SE = 0.086) and neutrality toward "Nuclear weapons contribute to peace by preventing conflict between countries" (mean = 0.080; SE = 0.090). Interestingly, these mean attitudes remained largely fixed across all interventions. We observed only weakly significant effects: the group and institution cues negatively affected responses to "Nuclear weapons help to keep my country safe" (respectively ATE = −0.216, $p = .032$; ATE = −0.218, $p = .033$). The group cue also had a weakly significant effect on fears about cheating that violates nuclear treaties (ATE = −0.132, $p = .040$). However, the significance of these results disappeared after we applied a Bonferroni-Holm correction due to the number of outcomes considered. The public may be more confident in its views on nuclear weapons and disarmament than on the specific approach taken by the TPNW. Disarmament advocacy has been around as long as the weapons themselves, but the TPNW's legal prohibition is distinct from the traditional phased warhead reductions of nuclear arms control.

Subgroup analysis

Although statistical power prevents a complete evaluation of heterogeneous effects, we nonetheless observed demographic trends that researchers would do well to examine in the future. Overall, we only observed significant differences in support for joining the TPNW based on (five-point) ideology and partisanship (coded as three-point party identification). Since these results tracked closely on one another, we focus primarily on partisanship (see app. K). Considering the control group, table 1 displays baseline majority support for the ban irrespective of party identification. Democrats were more positive about joining the TPNW than Independents but not significantly (73.6% vs. 63.8%; $p = .234$).

While Republicans were least supportive, a surprising majority (53.8%, SE = 5.6) preferred to join. Among Democrats and Independents, the institution cue yielded the largest effects (respective unadjusted estimates: conditional average treatment effect [CATE] = −16.1, $p = .005$; CATE = −15.4, $p = .091$). Among Republicans, it was the security cue (CATE = −25.5, $p < .001$). Pooling across arms, Republicans showed an ATE of −21 percentage points, versus −9.2 for Democrats and −9 for Independents. This difference may simply be the product of the greater favorability Republicans exhibit toward nuclear weapons relative to Democrats (Kull et al. 2004; Russett 1990–91), thus making them more sympathetic to critiques of disarmament. Members of each party also responded heterogeneously to treatment, although differences were generally insignificant at conventional levels.

4. Our results remain significant after applying a Bonferroni-Holm correction: institution cue ($p < .001$), security cue ($p < .001$), norms cue ($p = .046$), group cue ($p = .046$).

Table 1. Estimated Treatment Effects by Party ID (%)

		Arm			
	Control Mean	Group	Security	Norms	Institution
Democrat mean	73.6	−5.2	−12.2	−2.8	−16.1
	(4.2)	(4.8)	(4.6)	(4.3)	(4.7)
Independent mean	63.8	−5.8	−12.7	−5.0	−15.4
	(6.9)	(6.9)	(7.6)	(6.8)	(9.0)
Republican mean	53.8	−14.7	−25.5	−20.0	−24.4
	(5.6)	(4.9)	(4.9)	(5.3)	(4.7)

Note. Point estimates of mean responses are presented for each treatment arm. Bootstrapped SEs, computed with 10,000 replicates, are presented in parentheses under each corresponding point estimate. All estimates are unadjusted.

CONCLUSION

Even in a hostile information environment, negative government messages may galvanize public opinion against a popular emerging norm like nuclear nonpossession. This is especially true when messages are crafted around institutional and security rhetoric, as these cues yield particularly strong opposition to the TPNW. Trickle-down social effects emanating from coverage of official positions are less effective. Still, they too can have similarly persuasive effects to government use of normative rhetoric. In terms of the academic debate over elite cues versus group cues (e.g., Guisinger and Saunders 2017; Kertzer and Zeitzoff 2017), we find that both affect public views on foreign policy. However, certain elite messages containing substantive content have more powerful persuasive effects than social group pressures.

Antinormative messaging may be amplified by multiple cues. Individuals and groups can be affected heterogeneously, with Democrats most persuaded by the institution cue and Republicans by the security cue. But overall, our analyses showed that attitudes are largely fixed toward nuclear weapons but not toward the new ban approach—irrespective of widespread support for disarmament. This underscores the importance of winning the information battle on emerging norms. Varied content maximizes odds of striking a chord with different audiences and converting recipients into norm opponents who may transmit group cues. It should be heartening to ban advocates that opinion among even Republicans is supportive of the TPNW. Nonetheless, our analyses reveal that all demographic groups are persuadable and may increase opposition when presented with negative government messages and group cues. It is thus in the interest of norm advocates and critics to note the theoretical foundations underlying successful antinormative messaging and the utility of public outreach targeting different nodes of the information pipeline. If the ban continues to gain international support, we should expect a natural push and pull in the information cycle. Events like diplomacy with North Korea and the abrogation of arms control agreements like the Intermediate-Range Nuclear Forces Treaty will also increase public awareness of nuclear issues. Such news may offer opportunities to tailor messaging for and against the ban.

ACKNOWLEDGMENTS

The authors are grateful to David M. Allison, Peter M. Aronow, Alexander Coppock, Stephanie Kang, Natasza Marrouch, Molly Offer-Westort, Frances Rosenbluth, Lawrence Rubin, and attendees at the 2019 American Political Science Association Annual Meeting for comments. They would also like to extend special thanks to Courtenay Conrad and three anonymous reviewers.

REFERENCES

Bell, Mark S., and Kai Quek. 2015. "Realism, Idealism, and American Public Opinion on Nuclear Disarmament." MIT Political Science Department Research paper no. 2015-5. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2610450 (accessed April 20, 2020).

Casey-Maslen, Stuart. 2019. *The Treaty on the Prohibition of Nuclear Weapons: A Commentary*. Oxford: Oxford University Press.

Ford, Christopher Ashley. 2018. "The Treaty on the Prohibition of Nuclear Weapons: A Well-Intentioned Mistake." Speech, University of Iceland, Reykjavík. https://www.state.gov/remarks-and-releases-bureau-of-international-security-and-nonproliferation/the-treaty-on-the-prohibition-of-nuclear-weapons-a-well-intentioned-mistake/ (accessed April 20, 2020).

Gibbons, Rebecca Davis. 2019. "Addressing the Nuclear Ban Treaty." *Washington Quarterly* 42 (1): 27–40.

Guisinger, Alexandra, and Elizabeth N. Saunders. 2017. "Mapping the Boundaries of Elite Cues: How Elites Shape Mass Opinion across International Issues." *International Studies Quarterly* 61 (2): 425–41.

Keaten, Jamey, and Mark Lewis. 2017. "Nobel Peace Prize Awarded to Anti-nuclear Campaign Group." *Chicago Tribune*, October 6. https://www.chicagotribune.com/nation-world/ct-nobel-peace-prize-ican-20171006-story.html (accessed April 20, 2020).

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.

Kull, Steven, Clay Ramsay, Stefan Subias, and Evan Lewis. 2004. "PIPA/Knowledge Networks Poll: Americans on WMD Proliferation." Program

on International Policy Attitudes and Knowledge Networks Report. https://www.lasg.org/WMDreport_04_15_04.pdf (accessed April 20, 2020).

Mekata, Motoko. 2018. "How Transnational Civil Society Realized the Ban Treaty: An Interview with Beatrice Fihn." *Journal for Peace and Nuclear Disarmament* 1 (1): 79–92.

Milner, Helen V., and Dustin Tingley. 2015. *Sailing the Water's Edge: The Domestic Politics of American Foreign Policy*. Princeton, NJ: Princeton University Press.

Powlick, Philip J., and Andrew Z. Katz. 1998. "Defining the American Public Opinion/Foreign Policy Nexus." *Mershon International Studies Review* 42 (1): 29–61.

Russett, Bruce. 1990–91. "Doves, Hawks, and U.S. Public Opinion." *Political Science Quarterly* 105 (4): 515–38.

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.