

The Two Faces of Opposition to Chemical Weapons: Sincere Versus Insincere Norm-Holders*

Christopher W. Blair[†] Jonathan A. Chu[‡] and Joshua A. Schwartz[§]

Published in *Journal of Conflict Resolution*

Abstract

Prominent research holds that the use of weapons of mass destruction is taboo. But how strong are these norms? Investigating this question among the mass public, we argue that some citizens actually support taboo policies in private but are unwilling to express counter-normative opinions openly due to fear of social sanction. These insincere norm-holders are difficult to identify empirically because they are observationally equivalent to sincere norm-holders in direct-question surveys. To overcome this challenge, we use a list design, which allows survey respondents to indirectly express sensitive opinions. The results from three list experiments show that between 10% and 17% of Americans falsify their preferences over chemical weapons use when asked directly. Our findings advance a specific debate on the strength of weapons taboos, while our conceptualization of insincere norm-holders and methodological application have broader implications for how scholars might think about and measure norms in international politics.

*This is one of several joint articles by the authors; the ordering of names follows a principle of rotation with equal authorship implied. Generous support for this research was provided by the Christopher H. Browne Center for International Politics at the University of Pennsylvania and the Stanford Laboratory for the Study of American Values. We thank Thomas Dolan, Erik Lin-Greenberg, Michael Horowitz, Reid Pauly, and participants at the 2019 Norms and Behavioral Change Conference for helpful comments and advice.

[†]Ph.D. Candidate, Political Science, University of Pennsylvania, cwblair@sas.upenn.edu

[‡]Assistant Professor, National University of Singapore, jonathan.chu88@gmail.com

[§]Grand Strategy, Security, and Statecraft Fellow, Harvard Kennedy School; PhD Candidate, University of Pennsylvania; josha@sas.upenn.edu

Seminal studies argue that the use of certain weapons—most notably weapons of mass destruction like chemical and nuclear weapons—is taboo (Price 1997; Tannenwald 1999). While norms outline standards of appropriate behavior, taboos are stronger and more deeply internalized than regular norms because they connote “unthinkingness”—violations are viewed as so morally abhorrent that they are not even considered (Tannenwald 1999: 436). Recent research, however, suggests important limits to the strength of these alleged “taboos.” Members of the public and elites may be willing to break a taboo if doing so offers military advantages (Press, Sagan, and Valentino 2013; Pauly 2018; Rathbun and Stein 2019) or preserves other core values like saving the lives of co-national soldiers (Dolan 2013; Sagan and Valentino 2017; Smetana and Vranka 2020). Importantly, even when people support a norm, they may be doing so for consequentialist or strategic reasons rather than for reasons relating to intrinsic moral outrage (Press, Sagan, and Valentino 2013).

In this article, we draw on the broader literature on norms to more fully develop an additional rationale for why people might voice support for an anti-weapons norm. Because norms are morally salient, individuals may fear that publicly expressing counter-normative beliefs will lead to social sanctioning or other negative consequences (Shannon 2000). As a result, some respondents who do not genuinely, *privately* support a norm may still *publicly* express support when asked directly. We call these “insincere” norm-holders.

Why is it important to distinguish sincere from insincere norm-holders? If scholars and policymakers treat sincere and insincere norm-holders as equivalent, we will overstate the true strength of norms. As other literatures show, failing to appreciate the difference between true believers and preference falsifiers has political consequences. For example, Kuran (1995) shows that revolutions are often *ex ante* surprising and unpredictable because it is difficult to observe the distribution of citizens who sincerely versus insincerely support an authoritarian regime since voicing opposition could lead to sanctioning. Opposition to female political candidates follows a similar logic (Streb et al. 2008; Schwartz and Blair 2020). Ultimately, understanding and identifying insincere norm-holders helps us to better understand the strength and robustness of norms (McKeown 2009).

Despite their importance, insincere norm-holders are difficult to identify. Recent work in the weapons taboo literature has sought to measure the existence of norms among the general public using survey methodology (Press, Sagan, and Valentino 2013; Sagan and Valentino 2017; Rathbun and Stein 2019; Carpenter and Montgomery 2020; Koch and Wells 2020) and item-response techniques (Girard 2021), but these established approaches cannot distinguish between the sincere and insincere types. When asked directly, both sincere and insincere norm-holders will appear to support a norm.

We overcome this methodological challenge by employing list experiments, a technique used to elicit truthful responses about sensitive topics (Blair and Imai 2012), such as vote buying (Gonzalez-Ocantos et al. 2012; Frye, Reuter, and Szakonyi 2019), sexual violence (Traunmüller, Kijewski, and Freitag 2019), and holding racist (Kuklinski, Cobb, and Gilens, 1997; Gilens, Sniderman, and Kuklinski 1998) or sexist (Streb et al. 2008; Burden, Ono, and Yamada 2017) attitudes. By comparing the results from a list experiment with those from a traditional, direct-question survey, we are able to compare the proportions of sincere and insincere norm-holders.

In our inquiry, we focus primarily on the mass public, contributing to the recent literature on individual-level norm adherence (e.g., Sagan and Valentino 2018) and psychological approaches to issues of war more generally (e.g., Dolan 2016; Holmes and Yarhi-Milo 2017; Kertzer, Renshon, and Yarhi-Milo 2019; Chu, Holmes, and Traven 2021). Substantively, we focus chiefly on chemical weapons (CW), since insincere norm-holding is only likely when there are strong social desirability costs associated with publicly opposing a norm. We present additional evidence on nuclear weapons for contrast, but argue the social desirability costs are likely less severe for openly supporting nuclear use. Because Presidents Donald Trump and Barack Obama have both reaffirmed the CW taboo in recent years, the production and use of CW are illegal under international law, and the U.S. is in the process of completely destroying its CW arsenal (over 90% has been destroyed to date), the expected social costs of publicly supporting CW are likely relatively large (Kimball 2018). By contrast, the U.S. maintains a robust nuclear arsenal and a policy of nuclear deterrence. As such, and given the U.S.'s historical use

of nuclear weapons, the American public has a long tradition of openly discussing the merits of developing and deploying nuclear weapons, while there is no analogous public debate over CW.¹

Our results, which draw from three surveys of the U.S. public, show that insincere norm-holders exist and accounting for them nearly doubles the percentage of Americans that are willing to use CW in war. Support for CW is low overall, corroborating the existence of a general and relatively strong norm against them. However, for a significant portion of the public—around 25%—using CW is far from “unthinkable.” Moreover, the percentage of Americans that falsify their preferences over CW is substantively large, as it is on par with the percentage of white Americans that feigned support for affirmative action in 1994 (Gilens, Sniderman, and Kuklinski 1998: 171-172), the percentage of the American public that concealed their opposition to a female president in 2006 (Streb et al. 2008: 81-82), and the average proportion of respondents that falsified support for authoritarian regimes across a range of studies (Blair, Coppock, and Moor 2020: 1308-1309). The CW norm may therefore not be as robust as initially believed.

After discussing our main findings, we probe the generalizability of our framework to other political actors and types of weapons. First, in two follow-up list experiments, we find no evidence for insincere norm-holding in the context of public support for the use of nuclear weapons. This null finding is consistent with recent evidence on the nuclear taboo, as Press, Sagan, and Valentino (2013) show that support for nuclear weapons use is more about strategic concerns than social norms. More broadly, the null is likely due to the fact that possessing nuclear weapons (e.g., for deterrence) and using them in certain situations (e.g., in order to save American lives) is considered a more socially acceptable position than supporting the possession and use of CW (Smetana and Vranka 2020), which are banned internationally and confer little strategic benefits versus conventional arms.

Second, we present qualitative evidence that insincere norm-holding exists not just among the mass public, but also at the elite level. Most prominently, we argue that Syrian

¹Most recently, for example, Trump has publicly threatened to use nuclear weapons (Tannenwald 2018).

President Bashar al-Assad’s public, rhetorical support for the CW taboo belies his private opposition to the taboo demonstrated by his regime’s actions (Price 2019; Stimmer 2019). Although often described as the “poor man’s bomb” (Horowitz and Narang 2014), CW killed tens of thousands during World War I, and their increasing use by state and non-state actors from Russia and Syria to the Islamic State makes insincere norm-holding of the CW taboo—among elites and the mass public—especially policy relevant (Campbell and Murdie 2018).

Overall, this study contributes to two important literatures. First, extending the literature on norm robustness and regression (Shannon 2000; McKeown 2009; Panke and Petersohn 2012; Evers 2017; Deitelhoff and Zimmermann 2019; Stimmer 2019), we describe insincere norm-holding, explain how this shallow form of norm adherence can undermine the strength of norms, and offer a method to measure insincere norm-holding using survey experiments. Second, our results contribute to the literature on weapons taboos (Price 1997; Tannenwald 1999; Paul 2009; Dolan 2013; Press, Sagan, and Valentino 2013; Sagan and Valentino 2017; Pauly 2018; Rathbun and Stein 2019; Carpenter and Montgomery 2020; Koch and Wells 2020). In particular, we show that a significant portion of Americans falsify their support for the CW taboo, but not the nuclear taboo. Consequently, future research should take the possibility of insincere norm-holding into account when assessing the strength of norms and taboos. Some members of the public are two-faced in their support for norms.

Types of Norm-Holders

Traditional typologies of norm-holders consider whether individuals believe in a norm or not and whether believers support a norm due to the logic of consequences or the logic of appropriateness (March and Olsen 1998; Finnemore and Sikkink 1998). We build on these typologies by explicitly considering an additional factor: an actor’s private versus public view of a norm. Table 1 outlines a typology with two types of norm-holders—sincere and insincere—along with a third category of non norm-holders, or norm opponents. Sincere norm-holders publicly *and* privately support a norm due either to the

logic of consequences or the logic of appropriateness. For example, those who subscribe to the normative taboo against the use of CW, believing they are inherently immoral, are sincere norm-holders. Similarly, those opposing the use of CW for purely strategic reasons are still sincere in their beliefs, even if the underlying logic is one of consequences rather than appropriateness. Like sincere norm-holders, the public and private views of non norm-holders are also consistent. However, non norm-holders *do not support* a norm, either publicly or privately. Empirically, sincere and non norm-holders can be readily identified in traditional, direct-question surveys because they openly express their preferences.

Table 1: Types of Norm-Holders

	Public View of the Norm	Private View of the Norm	Detectable in Traditional Surveys?
Sincere Norm-Holders	Support	Support	Yes
Insincere Norm-Holders	Support	Do Not Support	No
Non Norm-Holders	Do Not Support	Do Not Support	Yes

By contrast, the public and private views of insincere norm-holders are inconsistent: they *publicly* support a norm but *privately* do not.² Consequently, insincere norm-holders are invisible in traditional surveys, as they will falsely report supporting a norm even though they oppose it in private.

Why would individuals be unwilling to express their true beliefs in surveys, or in other formats like interviews, public statements, or classified conversations? A large

²In principle, it is theoretically possible for insincere non norm-holders to exist. These would be individuals that publicly oppose a norm but privately support it. For example, in a discussion with an extremely hawkish friend, an individual could have incentives to support the use of normatively prohibited weapons like CW. Nevertheless, since by definition public support for a norm is much more likely to be socially desirable than opposition, we believe insincere norm-holding is much more likely than insincere non norm-holding. Our experimental results support this contention (or else direct support for the use of CW should have been higher than list support), but future research should consider the particular contexts in which insincere non norm-holding is most likely.

literature suggests that one of the principal explanations for preference falsification is social desirability bias.³ At least three factors must be present for social desirability bias to occur in the context of a particular issue (Blair, Coppock, and Moor 2020). First, a respondent must believe that one or more people (e.g., friends, family, co-workers, or researchers) or organizations (e.g., private firms or governments) hold a particular view on an issue. Second, a respondent must fear that their answer to a question about a sensitive issue will be observed by other people or organizations. Third, a respondent must conclude that failing to support the position held by observers will result in some cost. Typically, the costs are social and involve embarrassment, censure, or ostracization. For example, studies have shown that many respondents conceal racist or sexist beliefs out of fear of social sanction (Kuklinski, Cobb, and Gilens, 1997; Gilens, Sniderman, and Kuklinski 1998; Streb et al. 2008; Burden, Ono, and Yamada 2017). However, in extreme cases costs may involve monetary punishment like losing a job, electoral costs, or physical punishments (Blair, Coppock, and Moor 2020: 1299). Individuals that conceal their opposition to an authoritarian regime to avoid arrest, torture, or execution would fit into this latter category (Kuran 1995). The greater the expected costs of publicly supporting a certain position, the greater are the incentives for insincere norm-holding.

The general concept of preference falsification, or a disjuncture between public pronouncements and private beliefs and actions, is well-documented in previous literature (Kuran 1995; Krasner 1999; Lynch 2008; Finnemore 2009; Deitelhoff and Zimmermann 2019).⁴ Moreover, scholars of norms and taboos have argued that the types of costs that generate social desirability bias can also induce norm development and adherence (Finnemore and Sikkink 1998; Tannenwald 1999; Johnston 2001).

While building on this body of research, the contribution of our typology lies in the fact that existing categorizations of norms do not usually take into account insin-

³Blair, Coppock, and Moor (2020) argue that the term “sensitivity bias” is more precise. Note that insincere norm-holding differs from Wendt’s (1999) “first degree” of internalization because actors are not physically forced to support a norm, but merely face social pressures to do so.

⁴The psychology literature also makes a similar distinction between what we call sincere and insincere norm-holders. As Nail (1986: 195) describes, “conversion” entails public and private conformity (like sincere norm-holding), while “compliance” involves public conformity without private acceptance (like insincere norm-holding).

cere norm-holding explicitly as a separate category from sincere norm-holding based on a logic of consequences. For example, [Press, Sagan, and Valentino \(2013\)](#) and [Sagan and Valentino \(2017; 2018\)](#) distinguish between sincere norm-holders who support norms for moral reasons (logic of appropriateness) versus for strategic reasons, such as fearing setting a precedent for using WMDs (logic of consequences). In contrast, our study focuses on differentiating insincere from sincere norm-holders, rather than distinguishing among the different logics underpinning sincere norm-holding. As a result, our typology creates a theoretical framework that can guide future empirical research in detecting the full spectrum of norm-holders, which the current literature generally does not do.

By identifying insincere norm-holders, scholars can gain a better understanding of the public's true disposition toward various norms. Somewhat paradoxically, the presence of a significant number of insincere norm-holders may actually indicate a norm's relative strength rather than weakness. If people feel the need to publicly conform with opposition to CW, for example, then that suggests the norm is relatively well-developed ([Finnemore and Sikkink 1998](#)). If the norm was weaker, the costs of non-conformity would be lower, and a smaller fraction of the public would conceal their true opinions. From the perspective of norm proponents, insincere adherence to a norm is better than outright rejection. Whereas insincere norm-holders may adhere to norms due to fear of social sanction or accusations of hypocrisy ([Keck and Sikkink 1999](#)), non norm-holders are less susceptible to these pressures and may in fact relish transgression ([Evers 2017](#)). Overall, this means that the presence of insincere norm-holders, even in significant numbers, does *not* necessarily mean that a norm or taboo does not exist.

Nonetheless, as previously argued, the strongest norms are those that are sincerely held due to the logic of appropriateness rather than consequences ([Finnemore and Sikkink 1998](#); [Tannenwald 1999](#): 440; [Wendt 1999](#)). If individuals base their support for norms on a belief that violations are inherently wrong, then adherence is much more likely to endure challenges by norm opponents. In contrast, if many individuals privately oppose norms and only publicly support them due to social pressure, then the norm is ripe for regression: a shock to the social costliness of expressing opposition to the norm or acting in counter-

normative ways could rapidly lead these insincere norm holders to become outright norm opponents. Thus, failing to detect insincere norm-holding—which is a shallower form of norm-holding than sincere norm-holding—will lead scholars to overestimate the strength of norms.

Besides the costs of non-adherence, there are many other factors that affect the likelihood of insincere norm-holding. One relevant factor for this study relates to the context in which an individual is asked about a potentially sensitive opinion. The context most likely to elicit insincere norm-holding is a public forum like a signed petition or video interview posted online. In these settings, individuals must expose their views to a broad audience, heightening the risk that expressing a socially undesirable view will lead to social costs. The context least likely to provoke social desirability bias is an anonymous forum such as an online research survey, as the number of people or organizations that can observe (and attribute) a respondent’s view on an issue is limited. This makes our Internet-based survey experiments a hard test for identifying insincere norm-holders (Joinson 1999; Necka et al. 2016).⁵

This discussion highlights another important point about insincere norm-holding: it can vary across time and space. In one setting, an individual may feel the need to dissimulate their true views on a norm. In a different setting, the same individual may openly express their views. Changing strategic, political, or social circumstances over time can cause individuals to move between the different categories in Table 1.⁶ It is also important to note that outside our simplified framework, the intensity of support or opposition to a norm can vary. Some insincere norm-holders may strongly oppose a norm in private, while others may only do so hesitantly or with regret. Nonetheless, for a

⁵Private conversations among friends, family, co-workers, etc. likely fall somewhere in between. Incentives for insincere norm-holding in these conversations may be somewhat less than in the context of public forums because the number of people that can impose costs on a respondent for holding a particular view is more limited. Additionally, friends and family members are presumably more sympathetic to a respondent than strangers, perhaps reducing the costs of articulating a socially undesirable view. However, there can certainly be incentives for insincere norm-holding in these contexts. For example, in an analysis of classified discussions among U.S. government officials, Pauly (2018) found some evidence that reputational costs among peers was one factor discouraging officials from openly advocating nuclear use in wargames.

⁶This is especially the case for those whose support or opposition towards a norm is based on the logic of consequences. We elaborate on this point in the next section.

norm that involves an *absolute* prohibition on certain behavior—like how the CW taboo proscribes the use of CW in *any* circumstances—private willingness to violate the norm, even if only in rare scenarios, suggests that an individual cannot be described as a sincere norm-holder who privately supports the norm.

The Impact of Insincere Norm-Holders

Since insincere norm-holders falsify their preferences and are undetectable in traditional, direct-question surveys—and thus invisible in this context to political leaders and other members of the public—can they have any substantive political impact? Building on Kuran’s (1995) formal model and numerous empirical analyses (e.g., Gilens, Sniderman, and Kuklinski 1998; Cloward 2014), we argue that insincere norm-holders can have political impacts. This argument extends McKeown’s (2009) notion of the norm death cycle, the process by which a norm regresses. The status quo—in our case, an apparently strong taboo against CW—can obscure the relative fragility of a norm due to privately held opposition.

Consider latent opposition to a female president (Streb et al. 2008). Because many voters who privately oppose female candidates will not admit their preferences publicly in polls for fear of appearing sexist, they will not necessarily deter women from running for office. However, given that a woman decides to run, greater latent opposition will harm her at the ballot box, as sexist voters can anonymously express their opposition without fear of social sanction. A similar dynamic may operate with respect to CW use. Insincere CW norm-holders will not embolden politicians to *initiate* a challenge to the CW norm. But, if political decision-makers decide to challenge the norm, then greater latent counter-normative sentiment (e.g., greater private support for CW use) will make electoral punishment less likely, increasing the chances of further norm erosion.

The first step in the process of norm regression begins when a revisionist entrepreneur challenges a norm, either rhetorically, behaviorally, or both (Deitelhoff and Zimmermann 2019). Such challenges are often the result of significant security or economic shocks, like the September 11th attacks, which precipitated the use of torture by

U.S. forces, or the Syrian Civil War, which led to the use of CW by the Assad regime (Cortell and Peterson 1999; Kutz 2014). These shocks can change the cost-benefit analysis of norm transgression, thus incentivizing a challenge by political leaders.

Idiosyncratic leader preferences can also lead to norm-breaking, particularly if leaders are non or insincere norm-holders. For instance, President Trump publicly challenged the norm against the use of landmines in the absence of any significant security shock. Not all challenges are conducted in the same manner. Some may be carried out openly and candidly in the spirit of non norm-holding, while others may be denied completely or discounted rhetorically (e.g., “enhanced interrogation techniques”) in the spirit of insincere norm-holding.⁷

Whether a revisionist elite or other norm opponent can successfully challenge a norm depends on what McKeown (2009, 11) terms “challenge resonance,” or the receptivity of relevant audiences (e.g., the public) to the counter-normative position.⁸ The greater the latent support for overturning a norm among such an audience, perhaps due to a significant number of insincere norm-holders, the greater the probability of a reverse cascade, whereby the norm loses salience. Two reasons underpin this logic.

First, revisionist elites will face a lower penalty for challenging a norm since public backlash will be less likely. For example, in democracies, insincere norm-holders can anonymously express their genuine, counter-normative views at the ballot box without fear of sanction (Tomz, Weeks, and Yarhi-Milo 2020). The secret ballot therefore increases the chances that a revisionist entrepreneur will attain or remain in office, giving them an opportunity to continue to undermine a norm. Beyond the ballot box, a greater mass of people who privately do not support the norm means there will be less pushback overall, whether that takes the form of protests, general approval polls, public censure, etc. When revisionist leaders face limited or nonexistent backlash to their counter-normative behavior, they may infer the existence of insincere norm-holders, or at least learn their

⁷The stronger the norm is, the more likely challenges will be carried out by insincere norm-holders who violate a norm while *publicly* denying their violation in an attempt to reduce social, monetary, or physical costs.

⁸In other contexts, the relevant audience could be different. For example, the military would be a relevant audience in the case of a challenge to civil-military norms.

counter-normative challenges do not incur substantial costs. In these ways, opposition to a norm does not have to directly manifest in public opinion polls in order for insincere norm-holding to matter and be intuited by leaders.

Second, if an open challenge to a norm reduces the costs associated with publicly expressing opposition to it, then some insincere norm-holders may feel comfortable revealing their true opinions (Kuran 1995).⁹ For example, public support for torture and targeted killings by officials in President George W. Bush’s administration reduced the reputational costs of publicly supporting torture and assassination, despite previously prevailing norms against these practices (Kutz 2014).¹⁰ If insincere norm-holders become *public* opponents of a norm, then revisionists will be able to identify them in public polling and thus face reduced political pressure to abandon their challenge, making a reverse, counter-normative cascade more likely. Figure 1 summarizes this process and illustrates how latent opposition to the CW norm can have a political impact.

Insincere Norm-Holding and Chemical Weapons

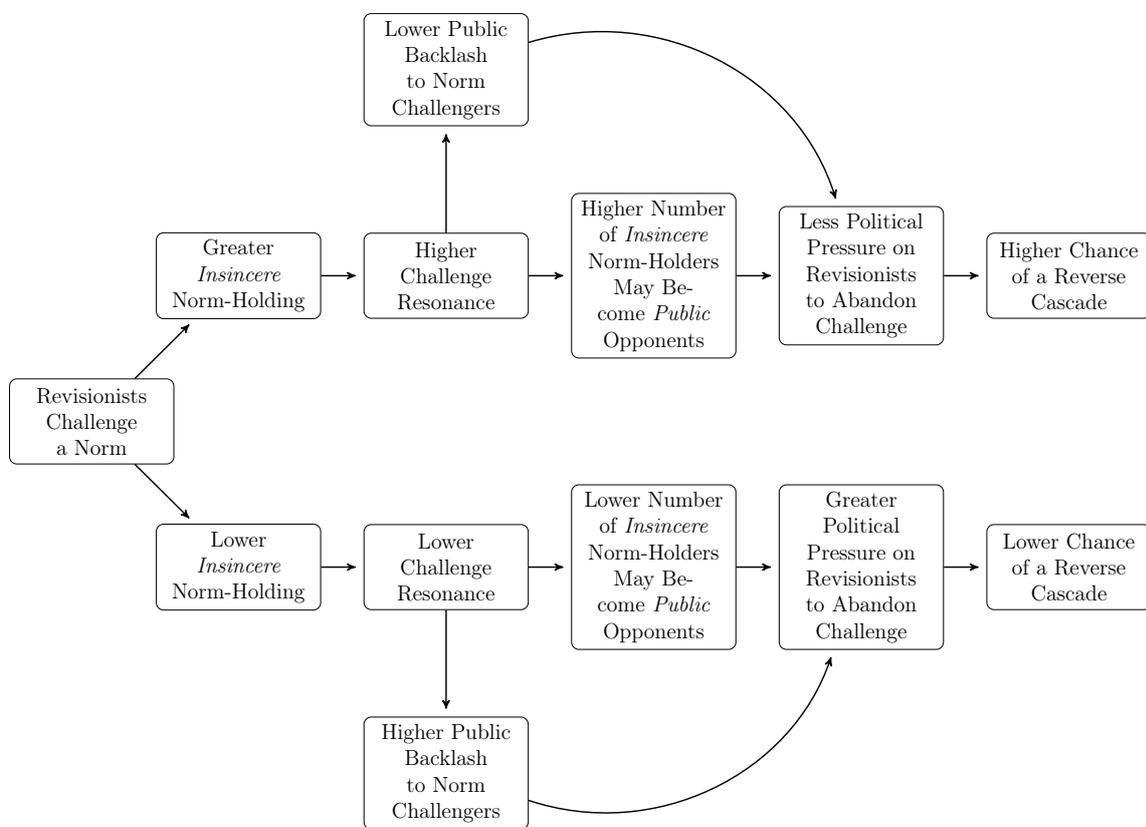
Insincere norm-holding is most likely to exist when there are strong social desirability costs associated with publicly opposing a norm. The norm or taboo against the use of CW is likely to fall under this category for several reasons.¹¹ First, the international legal prohibition against the use of CW is longstanding, well-developed, and nearly universal. Multi-country international agreements to ban the use of CW in war go back

⁹Kuran (1995) notes two individual-level benefits of honesty: “intrinsic utility” and “expressive utility.” Intrinsic utility is the benefit an individual derives from publicly supporting a privately-held belief, persuading others, and possibly tipping the scales in favor of their preferred alternative. Expressive utility is the benefit an individual receives merely by expressing their beliefs honestly, even if they cannot affect the outcome. If the costs of publicly expressing your private view of a norm go down due to an elite or other cue, then honesty should be more likely per Kuran’s (1995) framework due to the benefits of intrinsic and expressive utility.

¹⁰If norm violations are denied completely by the revisionist entrepreneur, then the costs of publicly expressing opposition to the norm may still remain relatively high.

¹¹Note that following Price (1997, 3), we define the chemical weapons taboo as a “norm proscribing the use of chemical weapons.” Therefore, while willingness to *enforce* a norm by exposing and punishing violators is relevant for norm robustness (e.g., Carnegie and Carson 2018; Price 2019), we do not utilize this factor to classify individuals as sincere, insincere, or non norm-holders of the CW norm in this study. For norms that explicitly involve enforcement—like the “responsibility to protect”—the public and private willingness of an actor to expose and punish violators would necessarily be used to categorize them as sincere, insincere, or non norm-holders.

Figure 1: How Insincere Norm-Holders Influence Norm Regression



Note: Figure adapted from models in Kuran (1995) and McKeown (2009).

at least as far as the 1800s, with the Brussels Convention in 1874 and the Hague Conference in 1899. Today, the most relevant agreement is the Chemical Weapons Convention, which fully bans the production, stockpiling, and use of CW, and has been signed and ratified by all but four countries in the world.¹² By contrast, the most comprehensive international nuclear weapons treaty—the Nuclear Nonproliferation Treaty—gives special status to “nuclear-weapon states” that manufactured and exploded nuclear devices before January 1, 1967. These countries are not required to immediately begin destroying nuclear weapons, while parties to the Chemical Weapons Convention must begin destruction within two years after ratification. Since previous research has demonstrated that international law can affect public opinion (Tomz 2008; Wallace 2013; Chu 2019), and the public is more likely to be informed about the legal status of these two types of WMD due to their prominence in the news,¹³ we expect that the stringent international legal prohibitions on CW should enhance the social desirability costs of openly opposing the CW norm.

Second, the U.S. government’s commitment to the Chemical Weapons Convention is not just rhetorical but also behavioral. To date, the U.S. has destroyed over 90% of its CW arsenal (Kimball 2018). By comparison, the U.S. maintains a large nuclear stockpile consisting of thousands of warheads, is currently spending hundreds of billions of dollars to modernize its nuclear arsenal, and remains publicly willing to use nuclear weapons in a range of circumstances, including in response to non-nuclear attacks.

Third, although the norm against the use of CW was challenged by the Assad regime (Deitelhoff and Zimmermann 2019), the *bipartisan* condemnation of this violation by both Democrats (e.g., President Obama) and Republicans (e.g., President Trump) served to reinforce this norm and increase the expected costs of openly opposing it among U.S. citizens by sending a strong elite cue. On the other hand, Trump has undermined norms surrounding nuclear weapons (Tannenwald 2018). In other contexts, the effect of

¹²North Korea, Egypt, and South Sudan have neither signed nor ratified. Israel has signed but not ratified.

¹³For example, that CW are banned under the Chemical Weapons Convention has been extensively discussed in the context of the Syrian case, and the public debate about the Treaty on the Prohibition of Nuclear Weapons has highlighted the more ambiguous legal status of nuclear weapons.

CW use might be different. For example, the use of deadly chemicals by the Russian government could undermine the CW norm and reduce the costs of openly opposing it among *Russian* citizens, who would be more exposed and susceptible to elite cues by the Russian government than U.S. citizens.

Finally, the prominent use of CW by the Syrian and Russian governments may have reinforced the CW norm and increased the expected costs of openly opposing it through a second mechanism: vivid information. Koch and Wells (2020) find evidence that support for nuclear use decreases when respondents are provided with vivid information about the human costs of a nuclear attack. Similarly, the use of CW in Syria and Russia, and the vivid, disturbing pictures and descriptions of its human impact (e.g. viral photos of attacks on Ghouta, Syria and Alexei Navalny in Russia) have likely served to enhance the social desirability costs associated with publicly opposing the norm (Geis and Schlag 2017).

Given these factors, we expect that U.S. citizens might reasonably believe that openly supporting the use of CW could lead to social costs like embarrassment, censure, or ostracization. This means that if some citizens privately oppose the norm against CW use, they would have incentives to publicly support it. Why might some citizens privately support the use of CW? For the same reasons they might support the use of other weapons: because they could provide tactical military advantages (Pauly 2018; Rathbun and Stein 2019); preserve other core values like saving the lives of co-national soldiers (Dolan 2013; Sagan and Valentino 2017; Smetana and Vranka 2020); enhance deterrence; maintain national sovereignty; or, in some people's minds, provide a potentially more painless method of incapacitating or killing enemy combatants (Bansak 2020). We therefore expect that insincere norm-holding will exist with respect to CW, and test the following observable implication of this argument:

H_1 : Some members of the public insincerely support the norm against chemical weapons use.¹⁴

¹⁴We also expect that a non-trivial percentage of American citizens will be insincere norm-holders, meaning our results will have substantive as well as statistical significance. If only, say, 3% of respondents turned out to be insincere norm-holders, then that would not be particularly noteworthy. However, following previous literature that utilizes list experiments, we hypothesize the existence of preference

Experimental Design

Identifying insincere norm-holders among the general public is difficult because they are observationally equivalent to sincere norm-holders in direct-question surveys. To overcome this challenge, we fielded two list experiments, which allowed respondents to indirectly express sensitive opinions on CW. The first experiment was conducted on a representative Qualtrics sample of 1,050 American adults in March 2016, and the second was fielded via Amazon mTurk on a convenience sample of 1,538 American adults in September/October 2019.¹⁵ Because the mTurk sample is non-representative, we weight responses in the second study to population demographics.¹⁶

In each list experiment, respondents were randomly assigned to one of three groups. The first group was directly asked whether they support using CW, the second group received a list of three policies, and the third group received an identical list but with one additional policy: the sensitive item—using CW during war.¹⁷ Respondents in the second and third groups were asked how many of the policies they support, but not which ones specifically. Consequently, respondents could express counter-normative positions anonymously. In accordance with previous literature, the average number of policies supported in the sensitive list (i.e., the list with four policies) minus the average number of policies supported in the list with only three policies equals the proportion of the sample that indirectly supports CW use. The difference between indirect support and direct support equals the proportion of respondents that are insincere norm-holders. In other words, the percentage of insincere norm-holders equals the percentage of respondents that support CW use when asked very discreetly (indirect support) minus the percentage of respondents that openly support CW use (direct support). Subjects that directly support

falsifiers rather than specifying an arbitrary proportion of respondents that we expect to falsify their true preferences.

¹⁵The second experiment was pre-registered with Evidence in Governance and Politics.

¹⁶Our weighting scheme was pre-registered. Unweighted results for study 2 are substantively similar. For more details, see the appendix.

¹⁷See the appendix for the full protocol of each experiment. In the mTurk study, we employ a quasi-combined list-direct design, where respondents that received the list of three policies without the question concerning CW also subsequently received the direct question (Aronow et al. 2015).

CW use are non norm-holders, and the rest of our sample are sincere norm-holders.

Experimental Results

Results from both experiments support our hypothesis that insincere norm-holders exist with respect to CW. Table 2 illustrates that support for using CW is higher when subjects are asked indirectly versus directly.¹⁸ Specifically, between 10% and 17% of Americans falsify their preferences over CW when asked directly. This is a substantively large effect, on par with the percentage of white Americans that falsified support for affirmative action in 1994 (Gilens, Sniderman, and Kuklinski 1998: 171-172), the percentage of the American public that concealed their opposition to a female president in 2006 (Streb et al. 2008: 81-82), and the average proportion of respondents that falsified support for authoritarian regimes across a range of studies (Blair, Coppock, and Moor 2020: 1308-1309). It is also greater than the average percentage of preferences falsifiers for vote buying, voter turnout, and racial, religious, or sexual orientation prejudice found in a meta-analysis of list experiments (Blair, Coppock, and Moor 2020: 1308-1309).

Table 2: Experimental Support for Chemical Weapons Use

	Indirect Support (%)	Direct Support (%)	Difference (Percentage Points)
Study 1	27.2%	16.6%	10.5*
Study 2 (Pre-Registered)	24.8%	8.5%	16.3**

Note: Results are calculated from 5,000 bootstraps. * = $p < 0.10$, ** = $p < 0.05$, where p-values indicate whether support for chemical weapons use is statistically greater than 0. Numbers may not sum due to rounding. The Qualtrics study was conducted in March 2016 and results are unweighted because it is a nationally representative sample. The mTurk study was conducted in September/October 2019, with results weighted to population demographics.

Further, taking insincere norm-holders into account suggests that more than one-quarter of the U.S. population would be permissive of using CW during war. The con-

¹⁸Following Coppock (2017), results are estimated using bootstrapped standard errors.

sistency of this result across our two studies—which took place over three years apart—builds confidence in our design, especially because the average number of policies supported in the list with only three policies is nearly identical across the two studies (1.46 in the Qualtrics study versus 1.45 in the mTurk study). The gap between indirect and direct support that we observe is also particularly striking given that our context—using two Internet-based survey experiments—is a hard test for social desirability since the surveys were anonymous (Joinson 1999; Necka et al., 2016). Finally, though the percentage of insincere norm-holders is not significantly different between studies 1 and 2 (p-value = 0.317), direct support is significantly lower in study 2 (p-value = 0.003). We cannot test precisely why this drop in direct support occurs, but the most likely explanation is that by striking Syria for CW use in April 2018 (between our surveys), Trump publicly reinforced the CW taboo and made clear support for the taboo was bipartisan, altering respondents’ social desirability calculus. In accordance with this argument, public polling shows that support for Trump’s military strike against Syria was higher than Obama’s proposed strike among Republicans and Independents, and remained steady among Democrats (Somin 2017).

In Table 3, we present the total percentages of each category of norm-holder—sincere, insincere, and non—across our experiments. These estimates represent a more accurate depiction of Americans’ views on CW than a traditional, direct-question survey would offer. Though a majority of U.S. citizens oppose the use of CW, corroborating the existence of a norm against them, a significant minority are insincere or non norm-holders.

Data from study 2 provides additional evidence supporting our proposed mechanism: social desirability. First, subjects that expressed support for CW in the direct question in study 2 were also asked to sign a public petition urging the U.S. government to adopt this position. About 43% of respondents that voiced support on the direct question still said they would not be comfortable signing a public petition. This finding suggests that many of the subjects that were willing to take a counter-normative position before our research team would still not be comfortable sharing their position with the broader public. This reinforces the notion that Internet-based surveys are a relatively

Table 3: The Distribution of Norm-Holders Across Our Experiments

	Sincere Norm-Holders (%)	Insincere Norm-Holders (%)	Non Norm-Holders (%)
Study 1	72.8%***	10.5%*	16.6%***
Study 2 (Pre-Registered)	75.2%***	16.3%**	8.5%***

Note: Results depict percentage support for chemical weapons use calculated from 5,000 bootstraps. * = $p < 0.10$, ** = $p < 0.05$, and *** = $p < 0.01$, where p-values indicate whether support for chemical weapons use is statistically greater than 0. Numbers may not sum due to rounding. The Qualtrics study was conducted in March 2016 and results are unweighted because it is a nationally representative sample. The mTurk study was conducted in September/October 2019, with results weighted to population demographics.

hard test for social desirability, and that insincere norm-holding is more likely in public forums. Second, all respondents in study 2 received an open-ended question about how their family and friends would react to the view that the U.S. should be willing to use CW. About 75% of respondents said their friends and family would react negatively to this position, and some of the most common words used were “horrified,” “appalled,” and “inhumane.” One respondent even articulated the possibility of social sanction if they expressed opposition to the taboo: “If I took that position I would be ostracized, and rightly so.”

Finally, to verify the robustness of our core results and further validate our design, we take a number of steps described in the appendix. First, substantively similar results emerge when we exclude respondents that failed the attention check. Second, we find no evidence of design effects using the test proposed by [Blair and Imai \(2012\)](#). Third, we model the determinants of direct support for CW use. Fourth, we explore multivariate list support for CW use using a maximum likelihood estimator ([Blair and Imai 2012](#)). Taken together, these tests lend confidence to our core finding that direct-question surveys substantially underestimate support for CW use.

Extensions

Our main contribution is to conceptualize insincere norm-holders and apply a methodology that allows us to empirically distinguish them. However, to unpack the broader implications of our argument, we explore two extensions. First, we examine whether our argument about insincere norm-holding extends to the nuclear taboo, a distinct norm from the CW non-use norm on which we focus. Second, to build further evidence in support of our CW results, we consider whether insincere norm-holding around CW exists among political elites, in addition to the mass public.

Application to the Nuclear Taboo

Does our argument about insincere norm-holding generalize to the nuclear taboo? As an extension, we apply our argument to assess whether insincere norm-holding also exists in the nuclear context. To do so, we replicated our study 2 mTurk design, replacing “be willing to use chemical weapons against other countries during war” with “be willing to use nuclear weapons against other countries during war.” With a total of 1,531 mTurk respondents, we find no evidence of insincere norm-holding with respect to nuclear weapons: 21.1% of the sample indirectly supports using nuclear weapons during war and 22.2% directly supports this policy.¹⁹

One possible explanation for this null result is that some respondents were thinking solely about nuclear deterrence; that is, using nuclear weapons to deter other nuclear-armed states. Since nuclear deterrence is a familiar, widely-accepted strategic concept and requires credibly threatening nuclear retaliation, some respondents may have been comfortable openly expressing their true beliefs, expecting few social costs no matter their opinion. If nuclear deterrence did truly drive this null finding, then we might have expected support for nuclear use to be higher than 22%. Nevertheless, to test this possibility, we fielded a third study in March 2020 on a representative sample of 1,204 Americans from Lucid.²⁰ We modified the direct question and sensitive list item

¹⁹This null result holds when we analyze the unweighted sample or exclude respondents that failed the attention check. See the appendix for a detailed description of the survey design and results.

²⁰See [Coppock and McClellan \(2019\)](#) for a description and validation of the Lucid sample and platform.

in this study to read: “Be willing to use nuclear weapons during war against countries that do not have nuclear weapons.” Since this policy has nothing to do with deterring nuclear-armed adversaries, we can test whether the null result from study 2 holds when considering nuclear use against non-nuclear foes, which may be seen as more socially undesirable. Again, we find no evidence of insincere norm-holding with respect to nuclear weapons: 16.8% of the sample indirectly supports using nuclear weapons and 19.6% directly supports this policy.²¹ When we exclude respondents that failed the attention check we find a small percentage of insincere norm-holders—approximately 3%—but this result is not statistically significant.

The robust null we find across these two studies ran counter to our initial expectations and is ripe for future research, but makes logical sense. Insincere norm-holding is only likely when there are strong social desirability costs associated with publicly opposing a norm. Given that Trump threatened to use nuclear weapons ([Tannenwald 2018](#)), the U.S. maintains a robust nuclear arsenal, and nuclear deterrence requires credibly threatening nuclear retaliation, the social costs for advocating willingness to use nuclear weapons in the abstract are likely lower than for supporting CW. For example, in a 2015 Pew Research poll 57% of respondents approved of the U.S.’s use of nuclear weapons against Japan in World War II ([Stokes 2015](#)). Although approval for this decision declined significantly from 1945 to 2015, the fact that over half of Americans are willing to openly support nuclear use against a non-nuclear-armed adversary suggests that the social desirability costs for advocating nuclear use in the abstract are relatively low. By contrast, in 1997 just 13% of Americans openly opposed the Chemical Weapons Convention ([Krepon, Smithson, and Parachini 1997](#)). Furthermore, our null result is consistent with [Smetana and Vranka’s \(2020\)](#) finding that support for using nuclear weapons in order to destroy an al-Qaeda WMD facility and save thousands of lives is significantly higher than support for using CW in the exact same scenario. Ultimately, Americans have a long tradition of openly debating the use of nuclear weapons, but not CW.

²¹This null result holds when we weight the sample to population demographics. See the appendix for a detailed description of the survey design and results.

Elite Insincere Norm-Holding Around Chemical Weapons

To apply our framework beyond the realm of public opinion, we also search for evidence of insincere norm-holding of the CW non-use norm among political elites. What evidence would establish that a leader is an insincere norm-holder and thus privately opposes the norm against the use of CW but publicly supports it? The most direct and unambiguous evidence that a political elite privately does not support the CW norm is when they purposely and repeatedly authorize the use of CW, as doing so reveals their true preferences. Beliefs and actions are, of course, highly interdependent, as beliefs affect actions and actions reveal beliefs. Since the CW taboo is absolute, the use of CW necessarily implies that an individual cannot be described as a sincere norm-holder; they do not privately believe in the norm against CW use. If a leader then proceeds to (a) publicly deny the use of CW and (b) rhetorically support the norm against CW, then they are an insincere norm-holder since they privately do not support the CW norm but publicly support it. In other words, when deeds do not match words, leaders reveal their hypocrisy and make clear they are insincere norm-holders. By contrast, leaders that employ CW and (a) do not deny their use or (b) argue that they are an acceptable weapon of war are non norm-holders. They privately *and* publicly oppose the CW norm.

Using this framework, it is clear that most, if not all, of the prominent cases involving the use of CW in recent years are associated with insincere norm-holding. Foremost among them is the Syrian government's repeated use of CW, which has been substantiated by reports from the [United Nations \(2016\)](#), [Organisation for the Prohibition of Chemical Weapons \(2020\)](#), and global think tanks. In fact, at the time of this writing, the Berlin-based Global Public Policy Institute has found credible evidence of about 350 attacks involving the use of CW in Syria, which suggests a significant degree of organization and planning ([Schneider and Lutkefend 2019](#)). Given the scale of these attacks and their persistence for almost a decade, it is highly likely that they are being conducted with the support of Syrian President Bashar al-Assad, who the U.S. intelligence community has assessed is "the ultimate decision maker for [Syria's] chemical weapons program"

(White House 2013).²² Consequently, at least since the security shock of the Syrian Civil War, Assad’s actions have revealed that he does not *privately* support the CW norm.²³ Yet, in accordance with the logic of insincere norm-holding and contrary to non norm-holding, he has also adamantly denied the use of CW and rhetorically supported the CW norm in *public*. For example, in one interview Assad claimed, “Even if we have [chemical weapons], we wouldn’t use them...We wouldn’t have the will, because morally this is not acceptable” (Agence France-Presse 2017). In another interview he said, “[The use of chemical weapons] is despicable. It is a crime” (Fox News 2013). Why would Assad rhetorically support the CW norm in public if, in private, he is willing to violate it? For the same reason that some individuals publicly support an authoritarian regime while privately opposing it: they fear the physical consequences of revealing their true opinions. Similarly, Assad worries that publicly opposing the CW norm could increase the chances of him losing his position or life due to an international military intervention and/or a further loss of domestic support.²⁴ Assad is therefore an insincere norm-holder: he publicly supports the norm against CW but does not support it in private. Overall, then, the logic of insincere norm-holding is quite similar for leaders like Assad and members of the public: both have incentives to avoid revealing their opposition to the CW norm because of the possibility of some type of punishment.

A similar dynamic holds for Russia. On March 4, 2018, in England, Russian operatives allegedly used a Novichok nerve agent in an assassination attempt against a former Russian double agent. They used a similar Novichok nerve agent on August 20, 2020, in order to poison Alexei Navalny, a political opponent of Vladimir Putin. A Russian FSB

²²Then Israeli defense minister Avigdor Lieberman also said he was “100 percent certain” that a 2017 chemical weapons attack was “directly ordered and planned by Assad” (Reuters 2017). While we believe it unlikely, we cannot completely rule out a scenario where Assad did not personally order Syria’s repeated CW attacks. If future evidence were to demonstrate this to be the case, then Assad could potentially be re-classified as a sincere norm-holder.

²³Note that while it is possible Assad did privately support the CW norm and thus was a sincere norm-holder before the Syrian Civil War produced military incentives for him to violate it, Syria’s pre-war production and stockpiling of CW combined with their refusal to sign the Chemical Weapons Convention implies that Assad may never have privately supported the CW taboo.

²⁴A similar logic has been used to explain why states privately engage in covert military interventions, but do not reveal their actions in public (e.g., Carson 2018; Poznansky 2019). Compared to our results among the U.S. public, which are likely driven by fear of social costs like embarrassment and censure, insincere norm-holding in this case is likely driven by physical costs.

agent, falsely believing he was speaking over the phone to a Russian National Security Council official, admitted to the attempted poisoning and explained that the nerve agent was applied to Navalny's underpants (Lister, Ward, and Shukla 2020). Nevertheless, Putin and the Russian government have *publicly* denied this allegation, Russia remains a member of the Chemical Weapons Convention, and Putin continues to hold that the use of CW is immoral (Price 2019).²⁵ Ultimately, Putin's public denials and support for the CW norm may help him avoid at least some domestic political costs, as one poll found *half* of Russians are skeptical Navalny was actually poisoned (Reuters 2020). Alternatively, if the Russia government had openly admitted their use of CW and lack of belief in the norm, it may have incurred greater domestic political costs.

Even the Islamic State has denied credible accusations that it used CW and rhetorically asserted their immorality (Price 2019). This despite their general willingness to publicly engage in other types of horrendous humanitarian offenses, such as recorded beheadings.

By contrast, in the past leaders and countries have been more candid about their use and support of CW, and thus have revealed themselves as non norm-holders. The most serious example of non norm-holding comes from World War I, where over 90,000 people were killed and more than 1,000,000 people were injured by the over 120,000 tons of poison gas used during the conflict (Thomas and Thomas 1970: 138; Spiers 1986: 13). By that time there was a nascent norm against the use of CW, as the Hague Convention of 1899 had outlawed "the use of projectiles the sole object of which is the diffusion of asphyxiating or deleterious gases" (Price 1997: 15). Still, and perhaps in a sign of things to come, Britain and the United States publicly opposed this measure. As Alfred Thayer Mahan, the famous naval strategist and U.S. representative at the conference, said,

"The use of projectiles of the kind in question can not be considered as being a means which is prohibited on the same ground as the poisoning of waters.

²⁵ Although Novichok was not added as a banned substance to the Chemical Weapons Convention until 2019 and the initial conception of the CW taboo proscribed the use of CW in warfare—not necessarily assassinations—Russia's use of CW in 2020 violated both the spirit and the letter of the Convention. The use of CW in targeted assassination attempts could provoke different reactions among domestic and international audiences than the use of CW in interstate or intrastate warfare, which could be an interesting avenue for future research.

Such projectiles might even be considered as more humane than those which kill or cripple in a much more cruel manner, by tearing the body with pieces of metal. Supposing that projectiles of this kind should be invented, their use may produce decisive results. Moreover, it would involve neither useless cruelty nor bad faith, as exists in the case of poisoning waters... [T]he use of those projectiles ought therefore to be considered as a lawful means of waging war” (Price 1997: 32).

Although the use of CW during the war was rampant, Germany was the first country to use lethal CW on a large scale. On April 22, 1915, at the Second Battle of Ypres in Belgium, Germany released chlorine gas from 6,000 cylinders, allowing a “favorable” wind to carry it towards unsuspecting French soldiers (Thomas and Thomas 1970: 138). The attack worked better than the Germans expected, as the French were driven from their positions and many choked to death. In accordance with non norm-holding, Germany did not deny that it had used CW and offered a variety of justifications for the deployment. For example, the German government argued that the use of cylinders was not a technical violation of the Hague Convention, which only outlawed the use of asphyxiating *projectiles*. German officials also asserted that CW were not more inhumane than other types of weapons (Price 1997: 51). According to a statement issued by Berlin,

“Anyone who has once seen the living picture of the hell provided by a rain of shells, hand grenades, subterranean mines and aerial bombs upon a bit of trench will certainly not regard slowly approaching clouds of smoke as a more inhumane method of warfare” (Garner 1920: 275).

That Assad, Putin, and most other violators of CW norm in today’s age are insincere norm-holders rather than non norm-holders—like Germany during World War I—is an important distinction for three reasons. First, insincere norm-holding suggests that the CW norm is relatively strong. If the norm was weaker, then Assad and others would feel free to openly admit and justify their use of CW, as Germany did, since the costs of doing so would be lower. Second, while the actions of modern CW norm violators constitute a behavioral challenge to the norm, their rhetorical condemnation of CW contributes to the norm’s robustness (Deitelhoff and Zimmermann 2019; Price 2019). On the other hand, non norm-holding challenges a norm both behaviorally *and* rhetorically, undermining the norm on both dimensions. Finally, the hypocrisy inherent in insincere norm-holding can

potentially be used to pressure norm transgressors to cease violations (Finnemore and Sikkink 1998; Risse, Ropp, and Sikkink 1999; Thomas 2001; Lynch 2008).

Of course, there is also a critical distinction between insincere and sincere norm-holders among elites. If Assad was instead a sincere norm-holder because he either truly believed CW were immoral or concluded it was strategically harmful to utilize them, then he would not have used CW in Syria. For example, prior to the invasion of Okinawa and Iwo Jima, Army Chief of Staff General George C. Marshall (who would later win the Nobel Peace Prize) and others recommended the use of CW (Moon 1989; Dolan 2013). Although doing so may have conferred military advantages and saved American lives, President Franklin Delano Roosevelt was a sincere holder of the CW norm and rejected all such proposals (Lovell 1963: 75).

Overall, this discussion illustrates that insincere norm-holding also exists among elites, and points to the ways in which identifying this relatively shallow form of norm adherence matters.

Conclusion

A prominent debate in political science surrounds the strength of taboos and norms in international politics. In this study we develop the concept of insincere norm-holders: individuals who conceal their true, counter-normative beliefs when asked directly for reasons of social desirability. Employing a series of list experiments, we establish the existence of insincere norm-holders in the context of CW use. More than 25% of the U.S. population is privately willing to use CW, and that number could rise significantly if political figures, or even ordinary people, begin to express public support, or if the U.S. can gain significant strategic advantages by using CW.²⁶

²⁶For example, replicating Press, Sagan, and Valentino's (2013) design, Smetana and Vranka (2020) find that over 40% of respondents would be willing to use CW to destroy an al-Qaeda WMD facility if doing so would save between 50,000 and 70,000 lives. While this study provides a hard-test for the CW taboo since the benefits of using CW are extremely large, it does not account for insincere norm-holding and so may actually be an underestimate of true support for CW use in this context. Note also that while support for torture among the U.S. public before 9/11 was measured in one poll as just 32%, President Bush still chose to violate the norm against torture (International Committee of the Red Cross 1999). A future president could therefore make a similar calculation with respect to CW.

Establishing a benchmark measure of public attitudes about CW provides an important corrective because these weapons have significant real-world impacts but are generally understudied. Future research should also recognize insincere norm-holders and utilize methods, like list experiments, to counteract social desirability bias. Despite the existence of insincere norm-holders, the overall low support for CW use corroborates the existence of a norm against them. Furthermore, the fact that some respondents feel the need to publicly conform with opposition to these weapons suggests that the norm is relatively well developed (Finnemore and Sikkink 1998). Nevertheless, our results are in accordance with recent evidence (Press, Sagan, and Valentino 2013; Sagan and Valentino 2017) that weapons taboos may be less entrenched than previously thought. Some members of the public are two-faced when it comes to opposition to CW.

This paper also highlights a number of promising avenues for future research. First, scholars should apply our theoretical framework and methodology to examine whether members of the public also falsify their preferences with respect to other norms, like the norm against torture or the norm against the use of landmines. Both of these norms have seen recent rhetorical challenges from elites in the U.S. and other countries, so studying the extent of insincere norm-holding could offer insights into their trajectory. Second, future work should examine whether there are cross-national differences in insincere norm-holding over CW, perhaps due to variation in freedom of speech or societal militarism. Third, scholars should analyze under what conditions members of the public have greater or lesser incentives to falsify their preferences with respect to chemical or other types of weapons. For example, recent research suggests that vivid information about the effects of WMDs (Koch and Wells 2020) or priming international law and ethical norms (Carpenter and Montgomery 2020) reduces support for their use, which may also increase incentives for insincere norm-holding. On the other hand, if CW are depicted as non-lethal alternatives to more traditional, destructive weapons, then that might increase support for their use and reduce the incentives for preference falsification (Bansak 2020). The intended target of CW could also alter the incentives for or against insincere norm-holding depending on where that target falls on the hierarchy of victims (Tezcür and Horschig

2020). Fourth, given the increasing attention to polling decision-makers (e.g., Hafner-Burton, LeVeck, and Victor 2017; Yarhi-Milo, Kertzer, and Renshon 2018), scholars should study the extent to which elites recognize and manipulate insincere norm-holding. Fifth, what is the relationship between insincere norm-holding—especially among elites—and the probability of norm enforcement by third parties? If an actor violates a norm but expresses support for it in public, does that lessen the probability of domestic and international blowback by giving states a greater ability to plausibly deny they violated the norm at all? For example, as discussed previously, Putin’s denials that he used CW to target political opponents and public support for the CW norm may have helped him avoid at least some domestic political costs. This dynamic may be similar to how covert action can reduce backlash to military actions (Carson 2018; Poznansky 2019). Finally, given challenges to the nuclear weapons, landmine, and other norms by President Trump, as well as the broader challenge to democratic norms around the world, more research should focus on how insincere norm-holding could affect the success of these revisionist efforts.

References

- Agence France-Presse. 2017. “Transcript of exclusive AFP interview with Syria’s Assad.” <https://www.thepeninsulaqatar.com/article/13/04/2017/Transcript-of-exclusive-AFP-interview-with-Syria-s-Assad-1>.
- Aronow, Peter M., Alexander Coppock, Forrest W. Crawford, and Donald P. Green. 2015. “Combining List Experiment and Direct Question Estimates of Sensitive Behavior Prevalence.” *Journal of Survey Statistics and Methodology* 3(1): 43-66.
- Bansak, Kirk. 2020. “Comparative Causal Mediation and Relaxing the Assumption of No Mediator–Outcome Confounding: An Application to International Law and Audience Costs.” *Political Analysis* 28(2): 1-22.
- Blair, Graeme, and Kosuke Imai. 2012. “Statistical Analysis of List Experiments.” *Political Analysis* 20(1): 47-77.
- Blair, Graeme, Alexander Coppock, and Margaret Moor. 2020. “When to Worry About Sensitivity Bias: Evidence from 30 Years of List Experiments.” *American Political Science Review* 114(4): 1297-315.
- Burden, Barry C., Yoshikuni Ono, and Masahiro Yamada. 2017. “Reassessing Public Support for a Female President.” *The Journal of Politics* 79(3): 1073-78.

- Campbell, Blake, and Amanda Murdie. 2018. "Keep the Informants Talking: The Pursuit and Use of CBRN Weapons by Terrorist Organizations." *Studies in Conflict & Terrorism*.
- Carnegie, Allison, and Austin Carson. 2018. "The Spotlight's Harsh Glare: Rethinking Publicity and International Order." *International Organization* 72(3): 627-57.
- Carpenter, Charli, and Alexander H. Montgomery. 2020. "The Stopping Power of Norms: Saturation Bombing, Civilian Immunity, and U.S. Attitudes Toward the Laws of War." *International Security* 45(2): 140-169.
- Carson, Austin. 2018. *Secret Wars: Covert Conflict in International Politics*. Princeton, NJ: Princeton University Press.
- Chu, Jonathan A. 2019. "A Clash of Norms? How Reciprocity and International Humanitarian Law affect American Opinion on the Treatment of POWs." *Journal of Conflict Resolution* 63(5): 1140-1164.
- Chu, Jonathan A., Marcus Holmes, and David Traven. 2021. "Inferring Intentions from Consequences: How Moral Judgments Shape Citizen Perceptions of Wartime Conduct." *Journal of Experimental Political Science* 8(2): 203-207.
- Cloward, Karisa. 2014. "False Commitments: Local Misrepresentation and the International Norms Against Female Genital Mutilation and Early Marriage." *International Organization* 68(3): 495-526.
- Coppock, Alexander. 2017. "Did Shy Trump Supporters Bias the 2016 Polls? Evidence from a Nationally-representative List Experiment." *Statistics, Politics, and Policy* 8(1): 29-40.
- Coppock, Alexander, and Oliver A. McClellan. 2019. "Validating the Demographic, Political, Psychological, and Experimental Results Obtained from a New Source of Online Survey Respondents." *Research and Politics* 6(1): 1-14.
- Cortell, Andrew P., and Susan Peterson. 1999. "Explaining Domestic Institutional Change." *British Journal of Political Science* 29(1): 177-203.
- Deitelhoff, Nicole, and Lisbeth Zimmermann. 2019. "Norms Under Challenge: Unpacking the Dynamics of Norm Robustness." *Journal of Global Security Studies* 4(1): 2-16.
- Dolan, Thomas M. 2013. "Unthinkable and Tragic: The Psychology of Weapons Taboos in War." *International Organization* 67(1): 37-63.
- Dolan, Thomas M. 2016. "Go Big or Go Home? Positive Emotions and Responses to Wartime Success." *International Studies Quarterly* 60(2): 230-42.
- Evers, Miles M. 2017. "On Transgression." *International Studies Quarterly* 61(4): 786-94.

- Finnemore, Martha, and Kathryn Sikkink. 1998. "International Norm Dynamics and Political Change." *International Organization* 52(4): 887-917.
- Finnemore, Martha. 2009. "Legitimacy, Hypocrisy, and the Social Structure of Unipolarity: Why Being a Unipole Isn't All It's Cracked Up to Be." *International Organization* 61(1): 58-85.
- Fox News. 2013. "Defiant Assad Claims Government Did Not Use Chem Weapons, Vows to Abide by Agreement." <https://www.foxnews.com/politics/defiant-assad-claims-government-did-not-use-chem-weapons-vows-to-abide-by-agreement>.
- Frye, Timothy, Ora John Reuter, and David Szakonyi. 2019. "Hitting Them With Carrots: Voter Intimidation and Vote Buying in Russia." *British Journal of Political Science* 49(3): 857-81.
- Garner, James Wilford. 1920. *International Law and the World War*. London, UK: Longmans, Green, and Co.
- Geis, Anna, and Gabi Schlag. 2017. "The Facts Cannot Be Denied: Legitimacy, War and the Use of Chemical Weapons in Syria." *Global Discourse* 7(2-3): 285-303.
- Gilens, Martin, Paul M. Sniderman, and James Kuklinski. 1998. "Affirmative Action and the Politics of Realignment." *British Journal of Political Science* 28(1): 159-83.
- Girard, Tyler. 2021. "Reconciling the Theoretical and Empirical Study of International Norms: A New Approach to Measurement." *American Political Science Review* 115(1): 331-38.
- Gonzalez-Ocantos, Ezequiel, Chad Kiewiet de Jonge, Carlos Meléndez, Javier Osorio, and David W. Nickerson. 2012. "Vote Buying and Social Desirability Bias: Experimental Evidence from Nicaragua" *American Journal of Political Science* 56(1): 202-17.
- Hafner-Burton, Emilie M., Brad L. LeVeck, and David G. Victor. 2017. "No False Promises: How the Prospect of Non-Compliance Affects Elite Preferences for International Cooperation." *International Studies Quarterly* 61(1): 136-49.
- Holmes, Marcus, and Keren Yarhi-Milo. 2017. "The Psychological Logic of Peace Summits: How Empathy Shapes Outcomes of Diplomatic Negotiations." *International Studies Quarterly* 61(1): 107-22.
- Horowitz, Michael C., and Neil Narang. 2014. "Poor Man's Atomic Bomb? Exploring the Relationship between Weapons of Mass Destruction." *Journal of Conflict Resolution* 58(3): 509-35.
- International Committee of the Red Cross. 1999. "International Committee of the Red Cross Poll: March 1999. Question 39: USGREEN.99PWAR.Q31." *Roper Center for Public Opinion Research*.

- Johnston, Alastair Iain. 2001. "Treating International Institutions as Social Environments." *International Studies Quarterly* 45(4): 487-515.
- Joinson, Adam. 1999. "Social Desirability, Anonymity, and Internet-based Questionnaires." *Behavior Research Methods, Instruments, & Computers* 31(3): 433-38.
- Keck, Margaret E., and Kathryn Sikkink. 1999. "Transnational Advocacy Networks in International and Regional Politics." *International Social Science Journal* 51(159): 89-101.
- Kertzer, Joshua D., Jonathan Renshon, and Keren Yarhi-Milo. 2019. "How Do Observers Assess Resolve?" *British Journal of Political Science* 51(1): 308-30.
- Kimball, Daryl. 2018. "Chemical Weapons: Frequently Asked Question." *Arms Control Association*.
- Koch, Lisa Langdon, and Matthew Wells. 2020. "Still taboo? Citizens' Attitudes Toward the Use of Nuclear Weapons." *Journal of Global Security Studies* 6(3): 1-18.
- Krasner, Stephen D. 1999. *Sovereignty: Organized Hypocrisy*. Princeton, NJ.: Princeton University Press.
- Krepon, Michael, Amy E. Smithson, and John Parachini. 1997. "The Battle to Obtain U.S. Ratification of the Chemical Weapons Convention." *The Henry L. Stimson Center, Occasional Paper No. 35*.
- Kuklinski, James H., Michael D. Cobb, and Martin Gilens. 1997. "Racial Attitudes and the New South." *The Journal of Politics* 59(2): 323-49.
- Kuran, Timur. 1995. *Private Truths, Public Lies: The Social Consequences of Preference Falsification*. Cambridge, MA: Harvard University Press.
- Kutz, Christopher. 2014. "How Norms Die: Torture and Assassination in American Security Policy." *Ethics & International Affairs* 28(4): 425-49.
- Lister, Tim, Clarissa Ward, and Sebastian Shukla. 2020. "Russian Opposition Leader Alexey Navalny Dupes Spy Into Revealing How He Was Poisoned." <https://www.cnn.com/2020/12/21/europe/russia-navalny-poisoning-underpants-ward/index.html>.
- Lovell, Stanley P. 1963. *Of Spies and Stratagems*. Englewood Cliffs, NJ: Prentice-Hall.
- Lynch, Marc. 2008. "Lie to Me: Sanctions on Iraq, Moral Argument and the International Politics of Hypocrisy." In *Moral Limit and Possibility in World Politics* edited by Richard M. Price, 165–196. Cambridge, UK: Cambridge University Press.
- March, James G., and Johan P. Olsen. 1998. "The Institutional Dynamics of International Political Orders." *International Organization* 52(4): 943-69.

- McKeown, Ryder. 2009. "Norm Regress: US Revisionism and the Slow Death of the Torture Norm." *International Relations* 23(1): 5-25.
- Moon, John Ellis van Courtland. 1989. "Chemical Warfare: A Forgotten Lesson." *Bulletin of the Atomic Scientists* 45(6): 40-43.
- Nail, Paul R. 1986. "Toward an Integration of Some Models and Theories of Social Response." *Psychological Bulletin* 100(2): 190-206.
- Necka, Elizabeth A., Stephanie Cacioppo, Greg J. Norman, and John T. Cacioppo. 2016. "Measuring the Prevalence of Problematic Respondent Behaviors among MTurk, Campus, and Community Participants." *PLoS ONE* 11(6): e0157732.
- Organisation for the Prohibition of Chemical Weapons. 2020. "First Report by the OPCW Investigation and Identification Team Pursuant to Paragraph 10 of Decision C-SS-4.Dec.3 'Addressing the Threat from Chemical Weapons Use' Ltamenah (Syrian Arab Republic) 24, 25, and 30 March 2017." <https://www.opcw.org/sites/default/files/documents/2020/04/s-1867-2020%28e%29.pdf>.
- Panke, Diana, and Ulrich Petersohn. 2012. "Why International Norms Disappear Sometimes." *European Journal of International Relations* 18(4): 719-42.
- Paul, T.V. 2009. *The Tradition of Non-Use of Nuclear Weapons*. Stanford, CA: Stanford University Press.
- 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.
- Poznansky, Michael. 2019. "Feigning Compliance: Covert Action and International Law." *International Studies Quarterly* 63(1): 72-84.
- 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.
- Price, Richard. 1995. "A Genealogy of the Chemical Weapons Taboo." *International Organization* 49(1): 73-103.
- Price, Richard. 1997. *The Chemical Weapons Taboo*. Ithaca, NY: Cornell University Press.
- Price, Richard. 2019. "Syria and the Chemical Weapons Taboo." *Journal of Global Security Studies* 4(1): 37-52.
- Rathbun, Brian C., and Rachel Stein. 2019. "Greater Goods: Morality and Attitudes toward the Use of Nuclear Weapons." *Journal of Conflict Resolution* 64(5): 787-816.

- Reuters. 2017. "Israel Says Assad's Forces Still Have Several Tonnes of Chemical Weapons." <https://www.reuters.com/article/us-mideast-crisis-syria-israel-idUSKBN17L2U8>.
- Reuters. 2020. "Half of Russians Sceptical Kremlin Critic Navalny Was Poisoned - Poll." <https://www.reuters.com/article/us-russia-politics-navalny/half-of-russians-sceptical-kremlin-critic-navalny-was-poisoned-poll-idUSKBN28Y1CX>.
- Risse, Thomas, Stephen Roppe, and Kathryn Sikkink. 1999. *The Power of Human Rights: International Norms and Domestic Change*. Cambridge, UK: Cambridge University Press.
- 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): 41-97.
- Sagan, Scott D., and Benjamin A. Valentino. 2018. "Not Just a War Theory: American Public Opinion on Ethics in Combat." *International Studies Quarterly* 62(3): 548-61.
- Schneider, Tobias, and Theresa Lutkefend. 2019. "Nowhere to Hide: The Logic of Chemical Weapons Use in Syria." *Global Public Policy Institute*.
- Schwartz, Joshua A., and Christopher W. Blair. 2020. "Do Women Make More Credible Threats? Gender Stereotypes, Audience Costs, and Crisis Bargaining." *International Organization* 74(4), 872-895.
- Shannon, Vaughn P. 2000. "Norms Are What States Make of Them: The Political Psychology of Norm Violation." *International Studies Quarterly* 44(2): 293-316.
- Smetana, Michael, and Marek Vranka. 2020. "How Moral Foundations Shape Public Approval of Nuclear, Chemical, and Conventional Strikes: New Evidence From Experimental Surveys." *International Interactions* 47(2): 374-90.
- Somin, Ilya. 2017. "Partisan Bias and Public Opinion on Trump's Syria Strike." <https://www.washingtonpost.com/news/volokh-conspiracy/wp/2017/04/13/48229/>.
- Spiers, Edward M. 1986. *Chemical Warfare*. New York, NY: Palgrave.
- Stimmer, Anette. 2019. "Beyond Internalization: Alternate Endings of the Norm Life Cycle." *International Studies Quarterly* 63(2): 270-80.
- Stokes Bruce. 2015. "70 Years After Hiroshima, Opinions Have Shifted on Use of Atomic Bomb." *Pew Research Center*.
- Streb, Matthew J., Barbara Burrell, Brian Frederick, and Michael A. Genovese. 2008. "Social Desirability Effects and Support for a Female President." *Public Opinion Quarterly* 72(1): 76-89.
- Tannenwald, Nina. 1999. "The Nuclear Taboo: The United States and the Normative

- Basis of Nuclear Non-Use.” *International Organization* 53(3): 433-68.
- Tannenwald, Nina. 2018. “How Strong Is the Nuclear Taboo Today?” *The Washington Quarterly* 41(3): 89-109.
- Tezcür, Güneş Murat, and Doreen Horschig. 2020. “A Conditional Norm: Chemical Warfare from Colonialism to Contemporary Civil Wars.” *Third World Quarterly* 42(2): 366-384.
- Thomas, Ann Van Wynen, and Aaron Joshua Thomas. 1970. *Legal Limits on the Use of Chemical and Biological Weapons*. University Park, TX: Southern Methodist University Press.
- Thomas, Daniel C. 2001. *The Helsinki Effect: International Norms, Human Rights, and the Demise of Communism*. Princeton, NJ: Princeton University Press.
- Tomz, Michael. 2008. “Reputation and the Effect of International Law on Preferences and Beliefs.” Unpublished manuscript, Department of Political Science, Stanford University.
- Tomz, Michael, Jessica L.P. Weeks, and Keren Yarhi-Milo. 2020. “Public Opinion and Decisions About Military Force in Democracies.” *International Organization* 74(1): 119-43.
- Trautmüller, Richard, Sara Kijewski, and Markus Freitag. 2019. “The Silent Victims of Sexual Violence during War: Evidence from a List Experiment in Sri Lanka.” *Journal of Conflict Resolution* 63(9): 2015-42.
- United Nations. 2016. “Security Council, S/2016/738, Third Report of the Organization for the Prohibition of Chemical Weapons–United Nations Joint Investigative Mechanism.” https://www.un.org/ga/search/view_doc.asp?symbol=S/2016/738.
- Wallace, Geoffrey P.R. 2013. “International Law and Public Attitudes Toward Torture: An Experimental Study.” *International Organization* 67(1): 105-140.
- Wendt, Alexander. 1999. *Social Theory of International Politics*. New York, NY: Cambridge University Press.
- White House. 2013. “Government Assessment of the Syrian Government’s Use of Chemical Weapons on August 21, 2013.”
- Yarhi-Milo, Keren, Joshua D. Kertzer, and Jonathan Renshon. 2018. “Tying Hands, Sinking Costs, and Leader Attributes.” *Journal of Conflict Resolution* 62(10): 2150-79.