Journal of Experimental Political Science (2026), **13**, 304–314
doi:10.1017/XPS.2025.10027

CAMBRIDGE UNIVERSITY PRESS

PREREGISTERED REPORT

A Replication and Extension of Willer et al. (2013), Overdoing Gender: A Test of the Masculine Overcompensation Thesis

Claire Gothreau[1] and Nicholas Haas[2]

[1]Rockefeller Center, Department of Government, Dartmouth College, Hanover, USA and [2]Department of Political Science, Aarhus University, Aarhus, Denmark
Corresponding author: Claire Gothreau; Email: Claire.M.Gothreau@dartmouth.edu

Abstract

Do men respond to a masculinity threat by adopting more conservative political attitudes? A highly cited 2013 study by Willer et al. – drawing on substantial work in social psychology – argues in the affirmative, reasoning that endorsing conservative views allows men to reaffirm their gender identity. In two experiments with student convenience samples (N_{total} 100–110, N_{men} 40–51), the authors find consistent evidence: inducing masculinity threat increases support for war, homophobic attitudes, and endorsement of dominance hierarchies. We conduct a preregistered replication of this foundational study using a nationally representative probability sample (N_{total} 2774, N_{men} 2073). Contrary to original findings, we observe no consistent evidence that masculinity threat alters political attitudes. We further do not find support for design differences between the replication and original study driving contrasting findings. Our results call into question the robustness of evidence linking masculinity threat to political attitudes and underscore the importance of re-evaluating widely accepted findings with representative, large samples.

Keywords: Masculinity; public opinion; replication; survey experiment; threat

In both popular culture and politics, many have asserted that there is a crisis of masculinity (Illing 2023; Kahloon 2023). Indeed, some politicians, such as U.S. Republican senator Josh Hawley – who has repeatedly called out liberals for their

This study received Institutional Review Board approval and was pre-registered with the Open Science Framework. The pre-registration document is available at https://doi.org/10.17605/OSF.IO/CSGFP. See Appendix D for information about the correspondence between the manuscript and the pre-registration document.

This article has earned badges for transparent research practices: Open data, Open materials, and Preregistered. For details see the Data Availability Statement.

© The Author(s), 2026. Published by Cambridge University Press on behalf of The American Political Science Association. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/XPS.2025.10027 Published online by Cambridge University Press

"attack on men" (Hawley 2021) – have gone so far as to make the issue a centerpiece of their campaign platforms. A crisis of masculinity is thought to partially undergird problems such as the global growth in far-right extremism, a movement that is chiefly driven by men and which often dovetails with backlash to gender equality, feminist movements, and the significant strides that women have made in achieving liberation (Greig 2019). Though much of the "crisis of masculinity" and its consequences may be attributed to structural factors like the decline in manufacturing, sociologists and psychologists have also attributed backlash to women's progress to a masculinity threat. Masculinity threat is the notion that manhood is a precarious state, and men must constantly affirm their gender, often through extreme demonstrations of masculinity (Dahl et al. 2015; Weaver and Vescio 2015). In an early and highly influential 2013 test of this theory published in the *American Journal of Sociology*, Willer and colleagues find that masculinity threat can impact political attitudes. Specifically, the authors observe in two lab experiments (N_{total} 100–110, N_{men} 40–51) conducted on a convenience sample of university students that inducing masculinity threat increases support for war, homophobic attitudes, and support for dominance hierarchies among male participants (Willer et al. 2013). We conduct a pre-registered replication of this foundational work with a large, nationally representative probability sample from the University of Chicago NORC AmeriSpeak Panel (N_{total} 2,774, N_{men} 2,095).[1]

Masculinity threat, also termed fragile masculinity or precarious manhood, refers to the theory that manhood is a status that is earned, maintained, and defended when it is challenged. Whereas womanhood is viewed as the result of natural biological development (Vandello and Bosson 2013), manhood and the masculine identity are understood to be more tenuous. Masculinity threat has been linked to maladaptive behaviors such as increased risk-taking (Parent et al. 2018) and physical aggression (Cohn et al. 2009), as well as to attitudinal consequences like decreased support for transgender rights (Harrison and Michelson 2019). In the pages that follow, we describe the original Willer et al. (2013) experiment, why the replication is of interest, the reasons it is unclear ex ante if the study will replicate, and our design and measurement strategy. Importantly, as we detail below, our design facilitates not only a replication of the study with a larger and more representative sample, but also insights into the *reasons* – theoretical or related to the treatment – behind the possible failure or success of the replication.

We find that Willer et al.'s (2013) main results do not hold in our replication: we do not observe that masculinity threat is associated with increased support for war, homophobic attitudes, support for dominance hierarchies, traditionalism, or other conservative attitudes. Our study falls somewhere between a direct and a conceptual replication of the original Willer et al. (2013) study, and differences in study design and implementation – in particular, time period, mode of treatment delivery, and sample characteristics – could be responsible for these contrasting findings. However, null findings do not appear to be driven by outdated outcome measures, respondent inattention or age, or the masculinity threat induction appearing

[1]More information on NORC and its collaboration with Time-sharing Experiments for the Social Sciences (TESS), which funded this project as part of its special competition on replications, can be found here.

https://doi.org/10.1017/XPS.2025.10027 Published online by Cambridge University Press

unrealistic; we also do not find any effects from inducing a more "general" sense of threat, and results indicate that we were sufficiently statistically powered to identify small effect sizes of interest. Although some influence of design differences cannot be completely ruled out, our analyses provide no evidence that they drive the contrasting findings. Further, if the original results are indeed specific to a particular sample, time period, or mode of delivery, such a finding could at the very least indicate important scope conditions for the underlying theory and its empirical test. Taken together, our results underscore the need for a more nuanced understanding of whether and when masculinity threat shapes political beliefs – and the types of men who may be most susceptible to such threats.

Original experiment and replication

Willer et al. (2013) conducted a between-subjects lab experiment to test the theory that masculinity threat impacts political attitudes. The goal of the experimental manipulation was to induce a sense of masculinity threat, which was predicted to lead to the adoption of more stereotypically masculine attitudes. In both Study 1 ($n = 111$) and Study 2 ($n = 100$),[2] the authors had university students fill out the Bem Sex Role Inventory (BSRI), a measure of stereotypically masculine and feminine traits. After completing the inventory, men and women were randomly assigned to receive feedback that they scored either in a masculine or feminine range. When men [women] received feedback that they scored in the feminine [masculine] range, this constituted the gender threat condition. When men [women] received feedback that they scored within the average masculine [feminine] range for their gender, this constituted the non-gender threat (control) condition. The dependent measures in Study 1 were three items on attitudes towards homosexuality, two items on support for the Iraq War, and a Car Purchasing Survey on vehicle desirability.[3] Study 2 dependent measures were eight dominance attitude items (Pratto et al. 1994), measures of political conservatism including attitudes on six different political items, eight system justification items (Jost et al. 2004), and seven items on traditionalism. After the post-test survey, respondents were debriefed about the deceptive gender identity feedback.

Our study both replicates the Willer et al. design and introduces innovations to further probe mechanisms and robustness.[4] First, we replicate the study with a much larger and nationally representative probability sample from the University of Chicago NORC AmeriSpeak Panel. The original Willer et al. study was conducted

[2]There are four studies in the Willer et al. (2013) paper. Our replication only concerns the first two, both of which were lab experiments. The study's central findings and test of the effects of masculinity threat on political attitudes come from these two studies. Study 3 is a correlational study, and Study 4 is a laboratory study involving the measurement of testosterone levels.

[3]Willer et al. also included other "filler" items in the political views survey.

[4]Space constraints precluded us from including *all* original study items. We therefore chose one of the Study 1 questions on the Iraq War, one on gay rights, and two on SUV purchasing. We supplement these measures with more updated ones described in the text. Similarly, we reduced Study 2 measures to include two social dominance items, two traditionalism items, one political conservatism item, and one system justification item. We also included a validated 15-item BSRI instead of the full 60-item validated battery from Willer et al. (2013).

https://doi.org/10.1017/XPS.2025.10027 Published online by Cambridge University Press

with a convenience sample of university students. It was also significantly underpowered and had a low number of men across both Study 1 and 2 – particularly problematic given the theoretical expectation and findings that gender threat only impacts men. Our minimum detectable effect size analyses indicate that we are sufficiently powered to detect effect sizes substantially smaller than those in the original study given our sample size of 2,774 respondents (see Appendix B). Importantly, because gender threat is theoretically a phenomenon specific to men, and empirical investigations into masculinity threat show that this is the case (see Carian and Sobotka 2018; Harrison and Mitchell 2019; DiMuccio and Knowles 2023), we oversample men (75% men, 24% women, 1% other).

A second contribution of our replication is the inclusion of a factual manipulation check asking respondents to recall their score on the experimentally manipulated feedback. We include this check to measure attention to the experimental stimuli without distorting treatment effects (Kane and Barabas 2019). Third, we introduce two additional conditions to the Willer et al. design. Given the lack of pre-test measures and piloting of the experimental manipulations, it is unclear if the treatment is inducing masculinity threat as opposed to a general sense of threat. Indeed, research indicates that people generally respond to threat with shifts toward conservative political and social positions (Bonanno and Jost 2006, but see Brandt et al. 2021). Participants randomly assigned to our "general threat" condition complete an entertainment and popular culture knowledge quiz and receive feedback that they performed poorly and have a low level of popular knowledge; importantly, feedback is not gendered. The selection of this domain is driven by the fact that research finds that women are stereotyped, by both women and men, as knowing more about this topic (Coffman 2014; Bordalo et al. 2019). Nevertheless, we expect that knowledge in this domain will be valued by both men and women, and therefore, negative feedback will induce a sense of threat. We view the condition as a way to distinguish whether a shift to more conservative attitudes among men truly stems from exposure to gender-specific threat or simply from threat evoked from receiving negative feedback in a valued domain.

The second condition we add to the design is a gender threat condition that is similar to the original treatment condition but constrains the range of feedback based on respondents' actual BSRI scores. One criticism of the original threat condition is that the feedback given may have seemed implausible to some participants. In particular, participants who answered the BSRI questions in alignment with gendered expectations – or, in other words, who had real scores that were on the extreme ends of the scale – may have felt the feedback was particularly implausible. Building on work by DiMuccio & Knowles (2023), we utilize participants' actual scores on the BSRI and subtract (or add) 20 points to their scores to induce gender threat.[5]

Fourth and finally, we supplement original outcome variables with updated measures, which better reflect current salient political issues. Additional questions probe opinions on transgender rights, support for decreasing the number of legal

[5]In the DiMuccio and Knowles (2023) paper, the authors only had conditions with the "constrained" feedback, and their dependent variables did not overlap with Willer et al. (2013). Their study thus did not constitute a replication.

https://doi.org/10.1017/XPS.2025.10027 Published online by Cambridge University Press

immigrants allowed to enter the United States, the desirability of electric vehicles, and preferential hiring policies for women. Finally, to establish discriminant validity, we include an item on marijuana legalization, which polling indicates is not strongly associated with gender or political leanings.

Data

We fielded our replication experiment on a nationally representative sample of American adults ($n = 2774$) through NORC's AmeriSpeak panel.[6] The AmeriSpeak panel is a probability sample with households selected from a sampling frame designed to provide a minimum of 97% coverage of the U.S. population. Because of our theoretical interest in gender threat towards men, we requested that they be oversampled; 74.7% of our sample identified as men and 24.2% identified as women (1.1% identified outside of the gender binary).[7] Including leaners, about 40.1% of our sample identified with the Democratic Party and 43.2% with the Republican Party (see Appendix Table C6 for full demographic information).

Replication of the original Willer et al. (2013) study

Table 1 displays replication results corresponding to Study 1 findings in Willer et al.; to facilitate comparison, we also mark statistically significant findings reported in the original study. Replicating their analysis strategy, we test hypotheses with a series of two-tailed *t*-tests comparing means across conditions split by binary gender identification (see Appendix Table C1 for robustness tests with *p*-values adjusted for multiple comparisons and Appendix Tables C8–C11 for regression models with demographic controls). All outcome variables are rescaled to range from 0 to 1 for interpretability and are coded such that higher values correspond to more conservative responses. In the original study, exposure to a gender identity threat increased men's support for war, homophobic attitudes, and interest in purchasing an SUV. In contrast, our replication finds no significant effect of the gender identity threat on any of these attitudes for either men or women.

Table 2 presents our replication results for the outcomes originally examined in Study 2 of Willer et al. The central finding in the original study was that gender identity threat led to increased endorsement of social dominance attitudes (Pratto et al. 1994) and traditionalist beliefs (Jost et al. 2003) among men. The authors also explored the effects of gender threat on system justification (Jost et al. 2004) and political conservatism. These results were interpreted as evidence that men respond to masculinity threats by reasserting dominance and reinforcing hierarchical beliefs. Again, we test hypotheses with a series of two-tailed *t*-tests comparing means across conditions, and we mark statistically significant findings from the Willer et al. study. Contrary to the original findings, we observe no significant effect of gender identity

[6]Our sample size for the replication of the original study treatment conditions – a comparison between gender identity threatened and not threatened conditions – is $N = 1{,}354$ ($N = 1{,}021$ men and $N = 333$ women; see Tables 1 and 2), with the remaining sample randomly allocated to either our constrained gender or general threat conditions.

[7]Respondents identifying in any categories other than "male" and "female" were randomly assigned to either the "feminine threat" or "masculine threat" condition.

https://doi.org/10.1017/XPS.2025.10027 Published online by Cambridge University Press

Table 1. *T*-test results replicating Willer et al. Study 1

Gender	Dependent variable	Gender identity threat (Mean)	No gender identity threat (Mean)	*T*-score and *P*-value
Women ($N = 333$)	Support for the Iraq War	0.36	0.37	$t = -0.11$, $p = 0.910$
	Negative view of gay rights	0.42	0.34	$t = 1.69$, $p = 0.092$
	SUV desirability	0.31	0.32	$t = -0.03$, $p = 0.972$
	SUV pay	0.07	0.09	$t = -1.65$, $p = 0.101$
Men ($N = 1{,}021$)	Support for the Iraq War	0.40	0.39	$t = 0.56$, $p = 0.574^{\dagger}$
	Negative view of gay rights	0.51	0.46	$t = 1.68$, $p = 0.093^{\dagger}$
	SUV desirability	0.33	0.37	$t = -1.43$, $p = 0.152^{\dagger}$
	SUV pay	0.08	0.09	$t = -1.56$, $p = 0.119^{\dagger}$

Notes: * $p < .05$, ** $p < .01$, *** $p < .001$. *T*-score and *p*-values correspond to our study; †indicates statistically significant difference ($p < 0.05$) in the original study.

threat on any of these outcomes for either men or women; again, results diverge from those in the original study.

Discussion

What explains our observed null results? One plausible reason for why our findings diverge could be that some of the original study's outcome measures – such as support for the Iraq War or opposition to gay rights – may no longer resonate with most Americans, as the Iraq War has ended and public opinion in the United States is now overwhelmingly in favor of same-sex marriage and broader gay rights (McCarthy 2022).[8] To evaluate if this shift in political context is responsible for our null findings, we can evaluate effects on our aforementioned supplementary items on salient contemporary issues on immigration, electric vehicles, transgender rights, and gender-based affirmative action policies.

Figure 1 summarizes results. The left-hand panel shows mean responses by experimental condition (control versus gender identity threat), while the right-hand panel displays average treatment effects. Once again, for women in our sample, we see no impact of gender identity threat on any of the contemporary political attitudes we measured. For men in our sample, we observe null effects for electric car desirability, support for legal immigration, marijuana legalization, and recognition of transgender

[8]Furthermore, it is possible that perceptions of the original attitudinal measures have changed. For example, negative views of gay people, support for the Iraq War, and SUV ownership may no longer be seen as masculine.

https://doi.org/10.1017/XPS.2025.10027 Published online by Cambridge University Press

Table 2. *T*-test results replicating Willer et al. Study 2

Gender	Dependent Variable	Gender identity threat (Mean)	No gender identity threat (Mean)	*T*-score and *P*-Value
Women ($N = 333$)	System justification	0.39	0.37	$t = 0.72$, $p = 0.471$
	Dominance attitudes	0.20	0.21	$t = -0.36$, $p = 0.722$
	Traditionalism	0.46	0.45	$t = 0.52$, $p = 0.602$
	Political conservatism	0.77	0.71	$t = 0.59$, $p = 0.554$
Men ($N = 1{,}021$))	System justification	0.39	0.39	$t = -0.2$, $p = 0.842$
	Dominance attitudes	0.24	0.25	$t = -0.53$, $p = 0.593^{\dagger}$
	Traditionalism	0.44	0.46	$t = -1.44$, $p = 0.150^{\dagger}$
	Political conservatism	0.87	0.88	$t = -0.26$, $p = 0.793$

Note: $p < .05$, ** $p < .01$, *** $p < .001$. *T*-score and *p*-values correspond to our study; † indicates statistically significant difference ($p < 0.05$) in original study.

identities. We do find that men whose gender identity is threatened are less supportive of gender-based affirmative action, as Willer et al. (2013)'s theory and other work linking masculinity threat to the assertion of dominance over women (Dahl et al. 2015) would suggest; however, this finding is not robust to correcting for multiple comparisons (see Appendix Table C1). In sum, we find little evidence that null results are attributable to outdated outcome measures.

A second explanation for our null finding is that some respondents may have questioned the credibility or realism of the treatment.[9] To examine this possibility, we can evaluate if treatment effects appear in our modified gender threat condition wherein participants received BSRI feedback constrained to not fall too far from their actual scores. A third explanation is that the original effects in Willer et al. were driven not by masculinity threat, but by a generally induced threat to which men were more responsive. If true, and the masculinity threat in our study did not induce a general threat – perhaps because masculinity was not sufficiently salient to respondents, or alternatively because it was so salient as to engender resistance to feedback – this could explain why our results differed from those in the original study. To investigate this explanation, we can see whether our general threat treatment condition produced effects.

Figure 2 summarizes results for these two additional treatments. For most of the original Willer et al. items, we find null effects for both our constrained masculinity and general threat treatments. Contrary to expectation, we observe that men in the

[9]One could further imagine that this concern was heightened in the replication as compared with the original study, as in the latter "participants were asked to wait several minutes" for their score, which was subsequently delivered to them in envelopes to which their name was affixed (Willer et al. 2013, p. 989).

https://doi.org/10.1017/XPS.2025.10027 Published online by Cambridge University Press

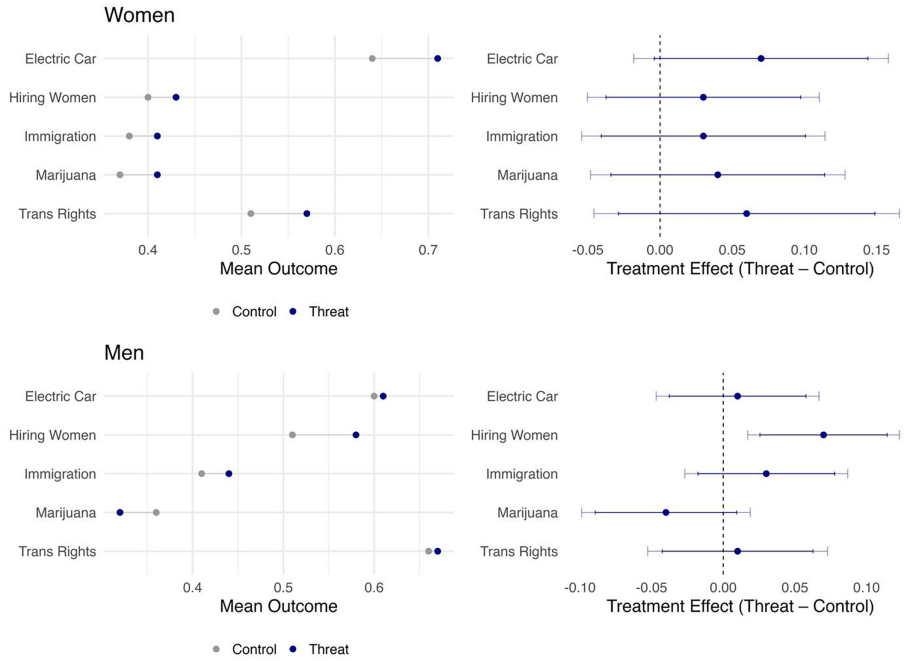


Figure 1. Means for updated outcome variables (left) and average treatment effects (right) split by binary gender.
Note: Darker bars indicate 95% confidence intervals, and thinner bars indicate 90% confidence intervals.

general threat condition report significantly lower social dominance attitudes. We also find that women in the general threat condition express significantly lower support for affirmative action policies and greater interest in purchasing an SUV. However, none of these effects remain statistically significant after correcting for multiple comparisons (see Appendix Table C1). Null results thus do not appear to be attributable to the credibility of our treatments or to the original study's findings being driven by general as opposed to masculinity threat.

In our Appendix, we rule out additional explanations for null results. Balance tests on observables confirm that randomization was effectively implemented (see Table C2); we appear sufficiently powered for our analyses (see Appendix B); differences in respondent attention do not appear to explain null results (see Tables C12–C15)[10]; and results are unchanged if we remove survey weights (see Table C16). Finally, we consider three further possible explanations: survey mode, as well as the size and characteristics of the original Willer et al. sample. Regarding mode, it is possible that effects in the original study are attributable to the treatment having greater salience in an in-person setting. We view this explanation as unlikely for several reasons. Our results do not differ between high- and low-attention respondents. We also find null results even in the more credible, constrained gender-threat condition. Moreover, online studies frequently detect effects from treatments that are considerably more complex than ours.

[10]Overall, 94% of participants passed the factual manipulation check, and 98.9% passed the attention check.

https://doi.org/10.1017/XPS.2025.10027 Published online by Cambridge University Press

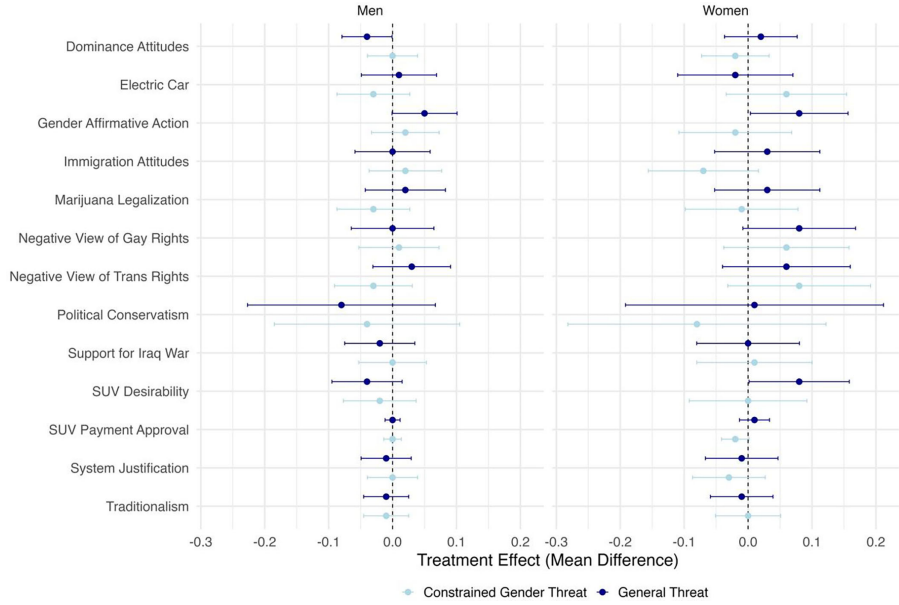


Figure 2. Average treatment effects by constrained feedback and general threat conditions.

Nevertheless, if this explanation were correct, it would imply that masculinity-threat effects arise only under a narrow and restrictive set of conditions.

As concerns sample characteristics, it could be the case either that masculinity threat is particularly relevant to the convenience sample of university students in the Willer et al. study or that it is especially easy to induce threat among this set of participants.[11] While we do not observe that age moderates treatment effects in our replication (see Appendix Table C21), we cannot rule out this explanation; however, as with mode effects, it would indicate a rather limited set of conditions under which we might expect induced threat to affect political attitudes. Finally, we note that the small sample size in the original study may have produced a type I error, or false positive, indicating either that experimentally inducing masculinity threat is unlikely to affect political attitudes or that the link between masculinity threat and such attitudes is not causal.

Conclusion

We conducted a replication and extension of foundational work by Willer et al. (2013) on the link between masculinity threat and political attitudes with a larger, nationally representative sample of U.S. adults. We do not find evidence that the original effects replicate. While we do not observe any evidence that contrasting findings are attributable to design differences between our replication and the original study, we cannot entirely rule out their potential impact. In particular, the move from a small,

[11]Notably, experimental findings by Harrison & Michelson (2019) also concern a small, convenience sample of university students.

https://doi.org/10.1017/XPS.2025.10027 Published online by Cambridge University Press

homogenous student sample in a controlled lab environment to a large, national probability sample is notable and could have altered the salience of the masculinity manipulation. Further, the two studies were conducted nearly twenty years apart; although we introduced updated dependent variables to capture possible changes in associations of different attitudes with masculinity, due to space constraints, we were unable to confirm whether these associations are realized in the U.S. public today.

Our results call for further examination into whether and when masculinity threat shapes political attitudes. Understanding which design differences, if any, are responsible for contrasting findings could further help to elucidate possible theoretical scope conditions. Are young men particularly likely to adopt conservative attitudes in response to masculinity threats? Do men assert their masculinity through different means or forms of expression today than they did 20 years ago? Are some settings particularly conducive for inducing masculinity threat? We conclude that more work needs to be done to investigate and clarify the relationship between masculinity threat and political attitudes.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/XPS.2025.10027.

Data availability. Support for this research was provided by Time-Sharing Experiments for the Social Sciences (TESS). Data collected by Time-sharing Experiments for the Social Sciences, NSF Grant 2017464, Maureen Craig, James Druckman, and Jeremy Freese, Principal Investigators. TESS is funded by the Social, Behavioral, and Economic Sciences Directorate of the National Science Foundation. The data, code, and any additional materials required to replicate all analyses in this article are available at the Journal of Experimental Political Science Dataverse within the Harvard Dataverse Network at https://doi.org/10.7910/DVN/5FVQBW.

Acknowledgements. In addition to the anonymous reviewers, we thank the members of the Center for the Experimental-Philosophical Study of Discrimination (CEPDISC, Aarhus University) for helpful comments on this project. We are also grateful to Robb Willer for his comments throughout the review process. We gratefully acknowledge funding support from Time-sharing Experiments for the Social Sciences (TESS) and the Danish National Research Foundation (DNRF144).

Competing interests. The authors are not aware of any conflicts of interest.

Ethics statement. This study received ethical approval from the Research Ethics Committee at Aarhus University (No. BSS-2024-048-S1). The research adheres to APSA's Principles and Guidance for Human Subjects Research. See Appendix E.

References

Bem, Sandra L. 1974. "The Measurement of Psychological Androgyny. *Journal of Consulting and Clinical Psychology 42*: 155–62.

Bonanno, George A and John T. Jost. 2006. "Conservative Shift Among High-Exposure Survivors of the September 11th Terrorist Attacks." *Basic and Applied Social Psychology 28*: 311–23.

Bordalo, Pedro, Katherine Coffman, Nicola Gennaioli and Andrei Shleifer. 2019. "Beliefs About Gender." *American Economic Review 109*: 739–73.

Brandt, Mark J, Felicity M. Turner-Zwinkels, Beste Karapirinler, Florian Van Leeuwen, Michael Bender, Yvette van Osch and Byron Adams. 2021. "The Association Between Threat and Politics Depends on the Type of Threat, the Political Domain, and the Country." *Personality And Social Psychology Bulletin 47*: 324–43.

Carian, Emily K and Tagart C. Sobotka. 2018. "Playing the Trump Card: Masculinity Threat and the U.S. 2016 Presidential Election." *Socius: Sociological Research for a Dynamic World 4*: 237802311774069. https://doi.org/10.1177/2378023117740699.

https://doi.org/10.1017/XPS.2025.10027 Published online by Cambridge University Press

Coffman, Katherine B. 2014. Evidence on Self-Stereotyping and the Contribution of Ideas. *The Quarterly Journal of Economics 129*: 1625–60. https://doi.org/10.1093/qje/qju023.

Cohn, Amy M, L. Alana Seibert and Amos Zeichner. 2009. "The Role of Restrictive Emotionality, Trait Anger, and Masculinity Threat in Men's Perpetration of Physical Aggression." *Psychology of Men & Masculinity 10*: 218–24. https://doi.org/10.1037/a0015151.

Dahl, Julia, Theresa Vescio and Kevin Weaver. 2015. "How Threats to Masculinity Sequentially Cause Public Discomfort, Anger, and Ideological Dominance over Women." *Social Psychology (Göttingen, Germany) 46*: 242–54. https://doi.org/10.1027/1864-9335/a000248.

DiMuccio, Sarah H and Knowles Eric D. 2023. "Something to Prove? Manhood Threats Increase Political Aggression Among Liberal Men." *Sex Roles 88:* 240–267. https://doi.org/10.1007/s11199-023-01349-x.

Gothreau, Claire and Nicholas Haas. 2025. Replication Data for: A Replication and Extension of Willer et al. (2013) Overdoing Gender: A Test of the Masculine Overcompensation Thesis, https://doi.org/10.7910/DVN/5FVQBW, Harvard Dataverse, DRAFT VERSION, UNF:6:uYZtrN6aEjSqYTNJEUdkzQ== [fileUNF] .

Greig, Alan. 2019. Masculinities and the Rise of the Far-Right. Implications for Oxfam's Work on Gender Justice.

Harrison, Brian and Melissa R Michelson. 2019. "Gender, Masculinity Threat, and Support for Transgender Rights: An Experimental Study." *Sex Roles 80*: 63–75. https://doi.org/10.1007/s11199-018-0916-6.

Hawley, Josh. 2021. "Senator Hawley Delivers National Conservatism Keynote on the Left's Attack on Men in America."https://www.hawley.senate.gov/senator-hawley-delivers-national-conservatism-keynote-lefts-attack-men-america.

Illing, Sean. 2023. "The New Crisis of Masculinity." 7 August. *Vox*. https://www.vox.com/the-gray-area/23813985/christine-emba-masculinity-the-gray-area.

Jost, John T, Mahzarin R Banaji and Brian A Nosek. 2004. "A Decade of System Justification Theory: Accumulated Evidence of Conscious and Unconscious Bolstering of the Status Quo." *Political Psychology 25*: 881–919.

Jost, John T, Jack Glaser, Arie W Kruglanski and Frank J Sulloway. 2003. "Political Conservatism as Motivated Social Cognition." *Psychological Bulletin 129*: 339–75. https://doi.org/10.1037/0033-2909.129.3.339.

Kahloon, Idrees. 2023. "What's the Matter with Men?." 23 January. *The New Yorker*. https://www.newyorker.com/magazine/2023/01/30/whats-the-matter-with-men.

Kane, John V and Jason Barabas. 2019. "No Harm in Checking: Using Factual Manipulation Checks to Assess Attentiveness in Experiments." *American Journal of Political Science 63*: 234–49.

McCarthy, Justin. 2022. "Same-Sex Marriage Support Inches Up to New High of 71%." 1 June. *Gallup*. https://news.gallup.com/poll/393197/same-sex-marriage-support-inches-new-high.aspx.

Parent, Mike C, Charlene M Kalenkoski and Eric Cardella. 2018. "Risky Business: Precarious Manhood and Investment Portfolio Decisions." *Psychology of Men & Masculinity 19*: 195.

Pratto, Felicia, Jim Sidanius, Lisa M Stallworth and Bertram F Malle. 1994. "Social Dominance Orientation: A Personality Variable Predicting Social and Political Attitudes." *Journal of Personality and Social Psychology 67*: 741–63.

Stanaland, Adam and Sarah Gaither. 2021. ""Be a Man": The Role of Social Pressure in Eliciting Men's Aggressive Cognition." *Personality & Social Psychology Bulletin 47*: 1596–611. https://doi.org/10.1177/0146167220984298.

Vandello, Joseph A and Jennifer K Bosson. 2013. "Hard Won and Easily Lost: A Review and Synthesis of Theory and Research on Precarious Manhood." *Psychology of Men & Masculinity 14*: 101–13. https://doi.org/10.1037/a0029826.

Weaver, Kevin S and Theresa K Vescio. 2015. "The Justification of Social Inequality in Response to Masculinity Threats." *Sex Roles 72*: 521–35. https://doi.org/10.1007/s11199-015-0484-y.

Willer, Robb, Bridget Conlon, Christabel L Rogalin and Michael T Wojnowicz. 2013. "Overdoing Gender: A Test of the Masculine Overcompensation Thesis." *The American Journal of Sociology 118*: 980–1022. https://doi.org/10.1086/668417.

Cite this article: Gothreau C and Haas N (2026). A Replication and Extension of Willer et al. (2013), Overdoing Gender: A Test of the Masculine Overcompensation Thesis. *Journal of Experimental Political Science* **13**, 304–314. https://doi.org/10.1017/XPS.2025.10027

https://doi.org/10.1017/XPS.2025.10027 Published online by Cambridge University Press