

Preregistered Direct Replication



Psychological Science 2021, Vol. 32(3) 451–458 © The Author(s) 2021 Article reuse guidelines: sagepub.com/journals-permissions DOI: 10.1177/0956797620968792 www.psychologicalscience.org/PS

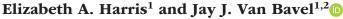












¹Department of Psychology, New York University, and ²Center for Neural Science, New York University

Abstract

There is currently a debate in political psychology about whether dogmatism and belief superiority are symmetric or asymmetric across the ideological spectrum. Toner, Leary, Asher, and Jongman-Sereno (2013) found that dogmatism was higher among conservatives than liberals, but both conservatives and liberals with extreme attitudes reported higher perceived superiority of beliefs. In the current study, we conducted a preregistered direct and conceptual replication of this previous research using a large nationally representative sample. Consistent with Toner et al.'s findings, our results showed that conservatives had higher dogmatism scores than liberals, whereas both conservative and liberal extreme attitudes were associated with higher belief superiority compared with more moderate attitudes. As in their study, we also found that whether conservative or liberal attitudes were associated with higher belief superiority was topic dependent. Contrasting Toner et al.'s findings, our results also showed that ideologically extreme individuals had higher dogmatism. We discuss implications of these results for theoretical debates in political psychology.

Keywords

replication, political psychology, belief superiority, dogmatism, ideology, open data, open materials, preregistered

Received 7/10/19; Revision accepted 8/28/20

In the United States, political polarization is on the rise, and this poses a significant threat to democracy (McCoy, Rahman, & Somer, 2018; Pew Research Center, 2014). Yet there is considerable debate about the origins and consequences of these political changes. Current research suggests that politically conservative individuals show higher levels of dogmatism, rigidity, and need for cognitive closure compared with liberal individuals (Jost, Glaser, Kruglanski, & Sulloway, 2003). However, a recent meta-analysis suggests that conservatives and liberals may both engage in similar levels of biased motivated reasoning (Ditto et al., 2019). It is relevant to note that a range of topics and measures are used in the literature, making it difficult to determine whether any similarities or differences are due to the specific issues under investigation (see Baron & Jost, 2019). The present research examined whether dogmatism and belief superiority are symmetric or asymmetric across the ideological spectrum using a range of measures and issues.

Toner, Leary, Asher, and Jongman-Sereno (2013) examined whether dogmatism and perceived superiority of one's beliefs is a partisan issue, that is, whether it is more prevalent among conservatives or found in members of both political orientations. Specifically, they asked liberals and conservatives to rate their attitudes toward controversial political issues and to report their perceived belief superiority on the topic, and they also measured participants' dogmatism. Belief superiority is conceptually distinct from general self-enhancement and resistance to attitude-inconsistent information (Hall & Raimi, 2018). Unlike belief certainty, belief superiority necessitates looking down on and thinking less of other people's positions (Raimi & Jongman-Sereno, 2020). Research on belief superiority suggests that it may be associated with a large bias in knowledge seeking (Hall

Corresponding Author:

Jay J. Van Bavel, New York University, Department of Psychology E-mail: jay.vanbavel@nyu.edu

& Raimi, 2018): People who report greater belief superiority tend to think of themselves as more knowledgeable, despite the fact that there is a greater gap between their perceived and actual knowledge. However, when belief superiority is lowered, people attend to information they previously regarded as inferior.

This research found that dogmatism was higher among conservatives than among liberals (asymmetry), but belief superiority was higher for both conservatives and liberals with extreme attitudes (symmetry; Toner et al., 2013). Moreover, whether the liberals or conservatives had higher perceived belief superiority varied from issue to issue. For example, individuals with very liberal attitudes regarding health care reported higher belief superiority than those with very conservative attitudes. Conversely, individuals with very conservative attitudes about affirmative action reported higher belief superiority than those with very liberal attitudes.

Given that assessing different attitudes and using certain measures can give rise to different findings in political psychology (e.g., Brandt, Reyna, Chambers, Crawford, & Wetherell, 2014), it is important to determine whether the primary findings are dependent on the specific operationalizations or aspects of the social and political context (see Van Bavel, Mende-Siedlecki, Brady, & Reinero, 2016). Therefore, we examined these issues systematically in our replication to determine whether prior findings were robust to these factors.

Toner and colleagues' (2013) original study appears to have been well executed, and the authors' findings appear likely to be replicated. However, the recent replication rate of psychology articles suggests that we should not take any outcome for granted (see Open Science Collaboration, 2015; Reinero et al., 2020). Therefore, we tested the generalizability of the original findings in multiple ways. First, we ensured that the results held when using the traditional one-item measure of political orientation. Unlike Toner and colleagues' inferred measure of political orientation, the one-item scale is widely used and has been shown to explain 85% of the variance of self-reported voting behavior (Jost, 2006). Second, we included the issue topic as a random effect in our belief-superiority analysis to generalize the results beyond the specific issues tested (see Yarkoni, 2019). Third, we collected a larger, more representative sample of participants. Much of psychology research using student samples (or Amazon's Mechanical Turk samples) is severely limited in generalizability (see Henrich, Heine, & Norenzayan, 2010). Therefore, collecting a representative sample is important for drawing conclusions that are generalizable to all Americans.

Method

We preregistered the following method details as well as our analysis plan on AsPredicted (https://aspredicted.org/fr299.pdf) prior to conducting any analyses. We also placed all materials, analysis scripts, and deidentified raw data on OSF (https://osf.io/x79pm). Our preregistration was designed to obtain a less biased effect-size estimate of the hypothesized relationships.

Participants

In their original study, Toner and colleagues (2013) used a large sample of 527 participants recruited on Amazon's Mechanical Turk. In our replication, we collected a larger sample of participants using the Lucid platform. Lucid uses quota-based sampling to ensure that participants are nationally representative in terms of age, gender, region, ethnicity, and political affiliation. Only participants who quit the study before reaching the end and those who failed the attention check were excluded.

According to a power analysis for a three-predictor multiple regression with a partial f^2 of .02 (a canonically small effect size) and an α of .01 (the α used in the original study), we needed 1,140 participants to achieve 95% power to detect an effect. Therefore, our target sample size was 1,300—considerably larger than in the original research. This specific power analysis was based on a linear multiple regression with belief superiority as the predicted variable and attitude as a linear term, attitude as a quadratic term, and dogmatism as predictors.

We also ran a power analysis for equivalence testing (specifically, regression-based two one-sided tests [TOST]) using the *TOSTER* (Version 0.3.4; Lakens, 2017) R package function powerTOSTone, in which we would be able to reject the presence of effects (ds) greater than 0.1. According to this power analysis, with an α of .05 and the proposed sample size of 1,140, we would have 92% power to detect an effect. Thus, our target sample size would be sufficient for equivalence tests.

Materials

We used the same materials as used in the original study, which the authors placed online in their Supplemental Material (Toner et al., 2013). These materials included a measure of dogmatism (specifically, Altemeyer's, 2002, dogmatism scale; $\alpha = .85$, M = 4.40, SD = 1.10), their measure of belief superiority (a 5-point Likert-type scale ranging from 1, no more correct than other viewpoints, to 5, totally correct, mine is the only correct view; $\alpha = .90$, M = 2.58, SD = 1.05), and their issue-relevant

attitudes ($\alpha = .58$, M = 2.66, SD = 0.60). These 4- or 5-point Likert-type scales include topics such as "When is it acceptable to torture?" (e.g., 1 = never, 2 = only inextreme circumstances to prevent an impending terrorist act, 3 = whenever it might yield useful information, 4 = all terrorists should be tortured). In addition to what was included in the original study, we also included items to assess political affiliation (self-identification as a Democrat, Republican, independent, or other) and political orientation (from 1, extremely conservative, to 7, extremely liberal; Jost, 2006). We also included oneitem measures of social and economic conservatism to tease apart the two forms of conservatism in our analyses: "In terms of [social and cultural issues/economic issues], how liberal or conservative are you?" (from 1, extremely conservative, to 7, extremely liberal).

We ran a separate pilot study prior to the replication. In this test, we asked participants whether they still considered the topics from the original study to be controversial issues (from 1, not at all controversial, to 7, extremely controversial). We also included new topics (lesbian, gay, bisexual, and transgender rights; Planned Parenthood funding; gender identity; death penalty; Confederate flag; euthanasia; safe spaces; cost of education; military spending; climate change; and gun control). Scales with a mean score above 4 (the midpoint of the scale) for both Democratic and Republican participants were interpreted as controversial and included in the main study. Additionally, we presented participants with the scale points for all the attitude scales (e.g., the torture attitude scale) and asked them to arrange them from extremely conservative to extremely liberal to confirm the validity of the scales.

Procedure

We used a procedure nearly identical to that used by Toner and colleagues (2013). As in the original study, participants were asked to fill out a questionnaire about controversial political issues (eight specific partisan issues, per the results of the pilot study). For each issue, participants stated their position on the corresponding 4- or 5-point scale and the perceived superiority of their belief relevant to that issue. Participants then completed Altemeyer's dogmatism measure. Finally, participants completed the demographics survey, including items about their age, gender, education level, to what extent they followed politics, their political affiliation, and their political orientation. We also included one-item measures of social and economic conservatism. Including these additional items after the primary dependent measures was one deviation from the original study.

Analysis plan

Dogmatism. Beginning with dogmatism, we used an analysis plan similar to that of Toner and colleagues (2013). We first analyzed whether there was a linear or quadratic relationship between political attitudes (i.e., the "inferred" political-orientation measure) and dogmatism to see whether individuals with more liberal or conservative political attitudes score more highly on dogmatism. We ran a multiple regression in which dogmatism was regressed onto a linear term for political attitudes and a quadratic term for political attitudes. Similarly, we ran the same analysis but with the predictor being self-reported political orientation, again with linear and quadratic terms. We also ran a simple t test comparing the average dogmatism scores of self-identified Democrats and Republicans. To infer support for the null hypothesis, if relevant, we planned to use equivalence testing for the regression predictors that were nonsignificant.

Across all analyses, we expected to successfully replicate the original study finding in three ways. First, the political-attitudes linear term would be a significant predictor of dogmatism; individuals with more conservative attitudes would show higher dogmatism than those with more liberal attitudes, and the quadratic term would be smaller than our smallest effect size of interest. Second, the political-orientation linear term would be a significant predictor; individuals with more conservative ideology would show higher dogmatism than those with more liberal ideology, and the quadratic term would be smaller than our smallest effect size of interest. Third, Republican, compared with Democratic, participants would report higher dogmatism. Only the first result was necessary for a successful exact replication (because this was the original measure used), but the second and third results examined generalizability.

If we replicated the finding that conservative individuals have higher trait dogmatism, we intended to run a final exploratory analysis. We would regress dogmatism onto participants' economic-conservatism score and social-conservatism score to examine which aspect of conservative ideology, or both, was correlated with dogmatism.

Belief superiority. For belief superiority, we planned to run a similar analysis to that of Toner and colleagues (2013), regressing belief superiority onto linear and quadratic effects for political attitudes. As they did, we planned to run the regressions both with and without including dogmatism as a covariate. Unlike in the original study, in which the authors ran a separate regression for each issue, our plan was to run one multilevel regression that included issue topic as a random variable and participant

as a random variable to generalize our results to controversial political issues more broadly and not just the specific topics we used. As in the dogmatism analysis, to infer support for the null hypothesis, if found, we planned to use equivalence testing for the regression predictors that we found to be nonsignificant. We predicted a successful replication, which we defined as the quadratic term, and not the linear term, being a significant predictor of belief superiority, indicating that both extreme liberal and extreme conservative attitudes would be associated with higher perceived belief superiority. We predicted that the linear term would be smaller than the smallest effect size of interest.

Results

Pilot study

Controversiality. We found that the overall mean score of controversiality for all 20 scales was above 4 (the midpoint of the scale). Looking at Democratic participants only, we found that all mean scores were above 4. Similarly, when we looked at only Republican participants, all mean scores were above 4. Finally, looking at just the independent participants, we found that all mean scores were above 4 except for the scales for government helping those in need and Muslim religious rights. Therefore, these two scales were not included in the replication study (per our preregistered exclusion criteria).

Scale-point ordering. We also wanted to include only scales for which at least 50% of the participants placed the scale points in their original order. These scales were (a) immigration, (b) abortion, (c) voting identification, (d) tax, (e) torturing terrorists, (f) affirmative action, (g) military spending, and (h) the government's handling of the COVID-19 pandemic. Of the original 20 scales included in the pilot study, these were the eight scales that met our criteria and were therefore included in the main study.

Main study

Participants. We collected data from 1,319 American participants on the platform Lucid. As preregistered, we removed participants who did not finish the study (n = 0) and those who failed either of the two attention checks (n = 612). Because of the large number of participants who failed the attention checks, we made a post hoc decision to run all analyses both with the smaller sample (n = 707) and with the total sample (N = 1,319). In our reduced sample, we had 707 participants. Politically, 41% of them self-categorized as Democrats and 32% as Republicans, 1 their mean age was 44.89 years (SD = 16.24), and 55% of them identified as female. Here, we report the

results from the smaller preregistered sample; the conclusions from the results of the larger sample were virtually identical (with the exception of the exploratory dogmatism analysis) and can be found at https://osf.io/x79pm.

Dogmatism. We sought to directly and conceptually replicate Toner et al.'s (2013) finding regarding dogmatism. They found that the linear term, but not the quadratic term, of average political attitude significantly predicted individuals' dogmatism score. First, we directly replicated this analysis. We found that both the linear, β = -0.39, t(703) = -5.68, p < .001, and the quadratic, β = 0.20, t(703) = 2.31, p = .021, mean-centered average attitude terms significantly predicted the average dogmatism score (see Fig. 1). This directly replicated the linear effect in the original work and revealed an additional curvilinear effect.

Second, we ran a conceptual-replication analysis, in which we regressed the mean dogmatism scores onto the linear and quadratic terms for mean-centered political orientation. We again found that both the linear, $\beta = -0.13$, t(702) = -5.21, p < .001, and the quadratic, $\beta = 0.06$, t(702) = 4.84, p < .001, political-orientation terms significantly predicted average dogmatism score (see Fig. 2). This conceptually replicated the linear effect in the original work and revealed an additional curvilinear effect.

Third, we ran another conceptual replication of the original findings, in which we conducted a simple t test comparing the mean dogmatism scores between self-identified Democrats and Republicans. The dogmatism scores of Democrats and Republicans were significantly different, t(468.74) = -4.81, p < .001, with Republicans having significantly higher scores.

Finally, we ran a preregistered exploratory analysis in which we regressed mean dogmatism scores onto self-reported social- and economic-conservatism scores. We found that both social conservatism, $\beta = -0.21$, t(703) = -5.40, p < .001, and economic conservatism, $\beta = 0.08$, t(703) = 2.09, p = .037, were significant predictors of dogmatism. However, when we ran this analysis with the total sample, only social conservatism, $\beta = -0.14$, t(1309) = -5.20, p < .001, and not economic conservatism, $\beta = 0.01$, t(1309) = 0.29, p = .77, was a significant predictor.

Belief superiority. As in the original study, we regressed belief superiority onto the linear and quadratic terms for attitude. Unlike in the original study, our plan was to run one multilevel model in which we included a random intercept for subject and a random intercept and linear and quadratic attitude terms for attitude topic. (We also ran the models without including the random quadratic term for attitude topic because of the boundary being

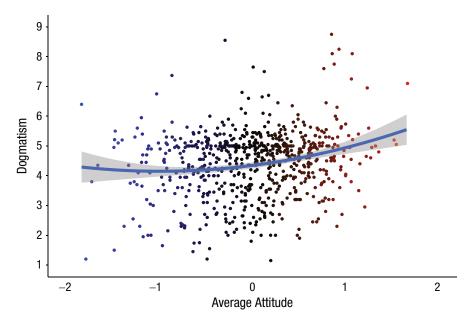


Fig. 1. Scatterplot showing the relation between dogmatism score (averaged across the 20-item measure) and attitude score (averaged across the eight attitude scales and then meancentered). The dogmatism scale runs from low (1) to high (9). The attitude scale runs from very liberal (low scores) to very conservative (high scores). The blue curve is the quadratic regression line, and the shaded region around it indicates the 95% confidence interval.

singular when included, and our conclusion about the results was the same; see https://osf.io/x79pm.) We ran this model with and without including dogmatism as a covariate. When we included dogmatism, we found that

the quadratic term, $\beta = 0.17$, t(54.02) = 18.69, p < .001, and dogmatism, $\beta = 0.41$, t(704.91) = 13.51, p < .001, but not the linear term, $\beta = -0.02$, t(6.86) = -0.27, p = .794, were significant predictors of belief superiority. When we

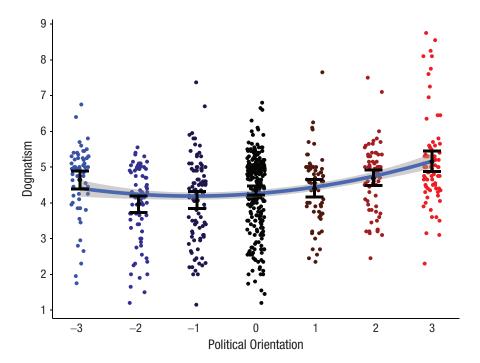


Fig. 2. Scatterplot showing the relation between dogmatism score (averaged across the 20-item measure) and political-orientation score (mean-centered). The dogmatism scale runs from low (1) to high (9). The political-orientation scale runs from very liberal (low scores) to very conservative (high scores). The blue curve is the quadratic regression line, and the shaded region around it indicates the 95% confidence interval. Within each group of political-orientation scores, the black dot with error bars represents the mean and associated 95% confidence interval.

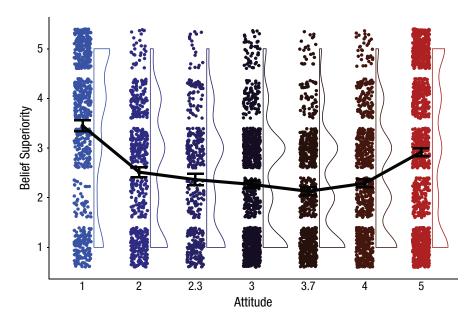


Fig. 3. Scatterplot showing the relation between belief-superiority rating (per subject and per topic) and the associated attitude rating. The belief-superiority scores run from low (1) to high (5). The attitude scores run from very liberal (1) to very conservative (5). Next to each set of raw data (dots) is a plot of the frequency of belief-superiority scores for that attitude value (the width of the plot indicates the density of the data). Within each group of attitude scores, the black dot with error bars represents the mean and associated 95% confidence interval.

did not include dogmatism, we again found that the quadratic term, $\beta = 0.17$, t(9.50) = 18.29, p < .001, but not the linear term, $\beta = -0.03$, t(6.85) = -0.37, p = .724, was a significant predictor of belief superiority (see Fig. 3).

We used equivalence testing to determine whether the nonsignificant linear terms were smaller than the smallest effect size of interest. We used the R package TOSTER and the function TOSTone. For the model including dogmatism, the equivalence test was significant, t(705) = 2.39, p = .0087, given equivalence bounds of -0.1 and 0.1 and an α of .05. The null-hypothesis test was nonsignificant, t(705) = -0.27, p = .786, given an α of .05. Similarly, for the model that did not include dogmatism, the equivalence test was significant, t(706) =2.29, p = .011, given equivalence bounds of -0.1 and 0.1 and an α of .05. The null-hypothesis test was nonsignificant, t(706) = -0.37, p = .713, given an α of .05. On the basis of the equivalence test and the nullhypothesis test combined, we can conclude that the observed linear effect is statistically not different from zero and statistically equivalent to zero.

Discussion

Toner and colleagues (2013) found that people with more conservative attitudes had higher dogmatism scores than those with more liberal attitudes and that people with more extreme political attitudes reported higher belief superiority. We successfully replicated this relationship between dogmatism and political attitudes in a large, nationally representative study. We also conceptually replicated this pattern using a measure of political orientation and political identity. We also successfully replicated the finding that individuals with more extreme attitudes, both liberal and conservative, reported higher belief superiority. Therefore, these patterns appear to be robust across samples, time, and measures.

Our research also revealed an important difference from prior work. Specifically, we found that individuals with more conservative or more liberal ideologies had higher dogmatism scores. It is hard to know whether this new finding stemmed from our larger, more representative sample or differences in the sociopolitical context that have changed over the past decade (see Van Bavel et al., 2016). For instance, this new result may be due to the increasing polarization among partisans (both Republicans and Democrats; Pew Research Center, 2014). It seems plausible, if not likely, that polarization could help produce greater dogmatism across the political spectrum.

The results speak to the debate on political symmetry versus asymmetry in political psychology (Baron & Jost, 2019; Ditto et al., 2019). Rather than resolving this debate, our findings appear to complicate the matter. We found clear and robust evidence for both symmetry (belief superiority and, to some extent, dogmatism) and

asymmetry (dogmatism) within the same sample of participants. We also ran exploratory analyses examining the relationship between belief superiority and attitude for each individual topic and found, as did Toner and colleagues (2013), that it was highly variable. For two of the topics (immigration and voting identification), more conservative attitudes were associated with higher belief superiority, whereas for the COVID-19 topic, more liberal attitudes were associated with higher belief superiority (military spending was trending in this direction as well; see https://osf.io/x79pm). Here, we selected our topics on the basis of those perceived as controversial, but future research should examine either a random sample of political topics or completely apolitical topics to help resolve this debate (and ideally include independent measures of ground truth). Until then, the symmetry-versus-asymmetry debate may represent a false dichotomy.

Conclusion

The current work provides further evidence that conservatives have higher dogmatism scores than liberals, whereas both conservative and liberal extreme attitudes are associated with higher belief superiority (and dogmatism). However, ideological differences in belief superiority vary by topic. Therefore, to assess general differences between liberals and conservatives, scholars must look across many diverse topics and model the data appropriately. If scholars instead choose to study one topic at a time, any ideological differences they find may say more about the topic than about innate differences between liberals and conservatives.

Transparency

Action Editor: D. Stephen Lindsay Editor: D. Stephen Lindsay

Author Contributions

E. A. Harris conceived of this replication study and developed the study materials with input from J. J. Van Bavel and three reviewers. E. A. Harris ran the analyses with input from J. J. Van Bavel. Both authors cowrote the manuscript and approved the final version for submission.

Declaration of Conflicting Interests

The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

Funding

This work was supported by a Social Sciences and Humanities Research Council of Canada Graduate Fellowship to E. A. Harris and a grant from the Alliance for Decision Education to J. J. Van Bavel.

Open Practices

Deidentified raw data, analysis scripts, and materials for both the main and pilot studies have been made publicly available via OSF and can be accessed at https://osf.io/x79pm. The design and analysis plans for this study were preregistered at https://aspredicted.org/fr299.pdf. This article has received the badges for Open Data, Open Materials, and Preregistration. More information about the Open Practices badges can be found at http://www.psy chologicalscience.org/publications/badges.







ORCID iD

Jay J. Van Bavel (b) https://orcid.org/0000-0002-2520-0442

Note

1. In a recent Gallup poll (Gallup, 2020), 30% of respondents reported being a Republican, and 31% reported being a Democrat. In our reduced sample, we had approximately the predicted percentage of Republicans (31.8%), although Democrats were somewhat oversampled (41.2%).

References

Altemeyer, B. (2002). Dogmatic behavior among students: Testing a new measure of dogmatism. *The Journal of Social Psychology*, *142*, 713–721. doi:10.1080/002245 40209603931

Baron, J., & Jost, J. T. (2019). False equivalence: Are liberals and conservatives in the United States equally biased? *Perspectives on Psychological Science*, *14*, 292–303.

Brandt, M. J., Reyna, C., Chambers, J. R., Crawford, J. T., & Wetherell, G. (2014). The ideological-conflict hypothesis: Intolerance among both liberals and conservatives. *Current Directions in Psychological Science*, *23*, 27–34.

Ditto, P. H., Liu, B. S., Clark, C. J., Wojcik, S. P., Chen, E. E., Grady, R. H., & Zinger, J. F. (2019). At least bias is bipartisan: A meta-analytic comparison of partisan bias in liberals and conservatives. *Perspectives on Psychological Science*, 14, 273–291.

Gallup. (2020). *Party affiliation*. Retrieved from https://news.gallup.com/poll/15370/party-affiliation.aspx

Hall, M. P., & Raimi, K. T. (2018). Is belief superiority justified by superior knowledge? *Journal of Experimental Social Psychology*, 76, 290–306.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, 33, 61–83.

Jost, J. T. (2006). The end of the end of ideology. *American Psychologist*, 61, 651–670.

Jost, J. T., Glaser, J., Kruglanski, A. W., & Sulloway, F. J. (2003). Political conservatism as motivated social cognition. *Psychological Bulletin*, 129, 339–375.

Lakens, D. (2017). Equivalence tests: A practical primer for *t* tests, correlations, and meta-analyses. *Social Psychological and Personality Science*, *8*, 355–362.

McCoy, J., Rahman, T., & Somer, M. (2018). Polarization and the global crisis of democracy: Common patterns, dynamics, and pernicious consequences for democratic politics. *American Behavioral Scientist*, *62*, 16–42.

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251), Article aac4716. doi:10.1126/science.aac4716

- Pew Research Center. (2014). *Political polarization in the American public*. Retrieved from www.people-press .org/2014/06/12/political-polarization-in-the-american-public/
- Raimi, K. T., & Jongman-Sereno, K. P. (2020). General belief superiority (GBS): Personality, motivation, and interpersonal relations. Self and Identity, 19, 546–571.
- Reinero, D. A., Wills, J. A., Brady, W. J., Mende-Siedlecki, P., Crawford, J. T., & Van Bavel, J. J. (2020). Is the political slant of psychology research related to scientific

- replicability? Perspectives on Psychological Science, 15, 1310–1328.
- Toner, K., Leary, M. R., Asher, M. W., & Jongman-Sereno, K. P. (2013). Feeling superior is a bipartisan issue: Extremity (not direction) of political views predicts perceived belief superiority. *Psychological Science*, *24*, 2454–2462. doi:10.1177/0956797613494848
- Van Bavel, J. J., Mende-Siedlecki, P., Brady, W. J., & Reinero, D. A. (2016). Contextual sensitivity in scientific reproducibility. *Proceedings of the National Academy of Sciences*, USA, 113, 6454–6459.
- Yarkoni, T. (2019). *The generalizability crisis*. PsyArXiv. doi:10.31234/osf.io/jqw35