Big if true: Fragility of the manipulable-voter model and the implications for politics and social science*

Andrew Gelman† John Bullock Gaurav Sood

4 Aug 2026

Abstract

The idea that voters are easily manipulated by seemingly minor stimuli has a long pedigree. But in the last two decades, a large set of studies has expanded the claim in ways that have not been sufficiently appreciated. Voters are now held to be affected not just by minor stimuli, but by stimuli that bear only very weak connections to the topic at hand. And increasingly, it is voters' behaviors, rather than only their survey responses, that are held to be affected. This "manipulable voter" model, if it were true, would have important implications for politics, supporting certain views held on the left, right, and technocratic center of the political spectrum. Conversely, the lack of empirical support for the model has implications for social science: if voters are not so easily swayed as the model suggests, we should reconsider the statistical and methodological practices that underpin studies that have been thought to support it.

1 Background

The claim that voters are easily manipulated, and manipulated by seemingly minor stimuli, has a long pedigree. For example, Kahneman and Tversky (1982) report that a seemingly meaningless change in wording caused 50% of their sample to change its preferences over policies that had been proposed to stop the spread of a deadly disease.[1] Bishop (2004) notes that support for the right of gay men and women to serve in the U.S. military rises by 10 percentage points when they are referred to as "gay" rather than "homosexual." And Zaller

*We thank Yotam Margalit for helpful comments.

†Department of Statistics and Department of Political Science, Columbia University.

[1]To our knowledge, none of the many efforts to replicate the experiment have found an effect as large as the 50% preference reversal reported by Kahneman and Tversky. But the weakest effect that we have found among replication efforts is a still-substantial 18% preference reversal.

(1992) devotes a book to explaining why small changes in question wording may cause people to change their answers to survey questions.

Findings like these have often been adduced as evidence that the voters are irrational or incompetent (e.g., Bishop 2004; Caplan 2011). There is an important sense in which such claims may be too strong. But even if we grant that, in some sense, these findings are evidence of irrationality, they arise from studies in which voters are behaving predictably, even reasonably. The effects in question are framing effects, and if voters are not responding to the content of the wording that is put to them, they are at least responding to the valence of that wording. For example, in the study that Bishop describes, it is not surprising that respondents were more supportive of gay rights when the critical word was "gay" rather than "homosexual": at the time of the study, "homosexual" was used as a pejorative term much more often than "gay." Still, it seems hard to deny that these framing studies point to a limitation of ordinary voters: the views that many of them express in surveys depend on small differences of wording that carry little or no substantive meaning.[2]

But in the last two decades, a large set of studies has expanded on the general claim of these framing studies in large and underappreciated ways. No longer is the claim only that people's views will be changed by small changes in issue framing or question wording. Instead, the claims are that behaviors as well as attitudes are subject to manipulation—and not just by the use of valenced wording, but through the work of stimuli that are seemingly irrelevant. Both of these claims are striking. The correspondence between attitudes and behaviors is famously weak (Greenwald 1968); in turn, it is one thing to claim that attitudes are changed by seemingly minor stimuli, and another thing—a much stronger thing—to claim that behaviors are also affected by minor stimuli. And it is again one thing to claim that people are affected by minor changes in stimuli that bear *some* obvious connection to the problem at hand—for example, the change from "gay" to "homosexual"—but quite another to claim that people are affected by stimuli that bear no obvious connection at all to the task at hand.

Many recent studies implicitly support one or both of these claims. For example:

- The claim of an electoral impact of shark attacks (Achen and Bartels 2002), disputed by Fowler and Hall (2018). The shark attack claim had to thread the needle: from one direction, it is essential for the argument that voting based on these events be irrational, as the larger story would fall apart if voters were legitimately retaliating for government failure. From the other direction, the research has to identify the vote changes specifically with shark attacks and not any correlated features.

- The claim that single women were 20 percentage points more likely to support Barack Obama when they were ovulating (Durante, Rae and Griskevicius 2013), disputed by Gelman (2015*b*).

- The claim that male college students with fat arms have more or less conservative attitudes on redistribution, depending on their parents' socioeconomic status (Petersen et al.

[2]The question of the extent to which people's survey responses reflect their actual views is beyond the scope of this essay, and we proceed as most others do, by taking people's statements about their own views at face value. See Bullock and Lenz (2019) for one examination of the question.

(2013), disputed by Gelman and Loken (2014)).

- The claim that flashing a smiley face on a computer screen for a fraction of a second causes large changes in attitudes on immigration (Bartels (2014), citing Erisen, Lodge and Taber (2014) and disputed by Gelman (2015*a*)).

- The claim that chance encounters with strangers on the street would have large and consistent effects on attitudes toward redistribution (Sands 2017). This particular claim was not well supported by the published data, which went in the opposite direction as hypothesized and could be explainable by chance, given the large variation in responses across people in this small study.

- The claim that exposure to a small picture of an American flag in the corner of a screen while one is taking a survey increases support for Republican candidates and views—up to eight months after the survey is taken (Carter, Ferguson and Hassin 2011).

These studies are often seen as being of a piece with the earlier literature on framing effects, in the sense that these newer studies, like the earlier ones, show that voters are irrational or easily manipulated. But the two major differences between these new studies and the earlier ones are at least as important as the similarity—and the differences have not been sufficiently appreciated. For brevity, we refer to the two new claims, taken together, as the *manipulable-voter model*.

This model, if true, has political implications that have not been much discussed.

On the other hand, if the general claims of the model are false or unsupported by the evidence, then, reconsideration of the studies that have given rise to the claims is in order. Did the authors of these studies draw faulty conclusions because of idiosyncratic errors, or are there common practices—the same set of faulty practices—that underpin much of this work?

In the next two sections, we take up each of these points in turn: first, the political implications of the model; second, the implications for social-science practice if the model is wrong. We then turn to an extended discussions a particular bodies of research involving experiments on voter turnout.

2 Political implications of the manipulable-voter model

What are the political implications of the model that attitudes and behavior are determined by normatively irrelevant stimuli, and a sort of neo-Freudian behavioral determinism, in which capricious voters are easily manipulated, slaves to hormones and to long-obsolete evolutionary influences, a world in which even the simple act of voting is labeled as "irrational," and with some gender essentialism (which might be called "schoolyard evolutionary biology") thrown in?

How to think about the model that voters are foolish and are being driven by some mixture of genes, hormones, and irrelevant stimuli? This is a story that fits the far left (democracy is a sham, no point in participating, better to have a revolution) and the center-right

(reduce the impact of politics on our lives, reduce taxes and business regulations so that decisions can be made by people with financial resources and skin in the game). That doesn't make the model descriptively wrong—for that, we have to point to statistical analyses and failed replications—but it can help to explain some of its appeal, especially for those who seek validation of traditional sex roles.

The model that voters are stupid (or, more charitably, ignorant, or, even more charitably, are rationally allocating their time to not bother becoming informed about their political choices) can also appeal to technocrats of the "Nudge" variety (Thaler and Sunstein (2008); see also (Szászi et al. 2022)): the flip side of the belief that attitudes and behaviors are predictably influenced by stimuli that should be irrelevant, is the belief that they can be deliberately manipulated by outsiders, whether they be malicious advertisers and political consultants, or benevolent paternalists. This general attitude also applies to the scholars of earlier research on framing, which is why we take pains to distinguish the strong form of the manipulable-voter model (in which there are large and consistent effects from essentially irrelevant stimuli) from that earlier literature which made more defensible claims.

Compare the above "behaviorist" attitude with traditional political science and sociology, which focus on decision making of political and social actors, power, coordination, roles, and so on. For example, in the area of public opinion and voting, consider Ferguson (1995); Uhlaner (1989); Hirano et al. (2010); Lax and Phillips (2009); Ashworth and De Mesquita (2014), which come from different perspectives and represent different schools and subfields of political science, but all with an overall picture of individuals and groups with clearly-defined aims and ideologies, jockeying for power in a somewhat rational way, with political and social mobilization serving to mediate the challenges of mass politics.

We consider this mainstream approach broadly—we are *not* restricting ourselves to rational-actor models, as much of political and social science recognizes that people have a mix of motivations and influences. Political actors make mistakes, even prospectively, and voters make incoherent choices and do not always have a clear understanding of the positions taken by candidates they vote for or against.

2.1 Connection between shortcut scientific methods and shortcut models of political attitudes and behavior

To get people to do something, you convince them it's a good idea or you remove obstacles in their way. Both these tasks can be difficult, hence we understand the appeal of shortcuts.

The push-a-button, take-a-pill model of social science is too simple in that it supposes that people can be consistently manipulated by small, sometimes barely perceptible, interventions; it is too complicated in that it attempts to sidestep more direct methods of changing attitudes and behavior that can require hard work. For example, views on redistribution can change based on personal or national economic conditions or the reframing of political issues such as trade and immigration. For some issues, focused canvassing can change opinions—but even after these very resource-intensive efforts, the effects after a few days are typically less than one-tenth of one standard deviation (Kalla and Broockman 2020). Compared to these di-

rect exposures or interventions, it would be surprising if an innocuous chance encounter with a stranger were to have a large and consistent effect. Again, if such effects were real, if we truly "had no choice but to accept that these conclusions are true"[3] then we would have to reassess our understanding of political attitudes and behavior.

We also see a resemblance between the push-a-button model of behavior and a shallow, automatic form of social science in which the "buttons" being pushed are randomized experiments, identification strategies in observational studies, and statistical significance. A lack of concern about details (as in the studies on ovulation and behavior that got the dates of ovulation wrong; see discussion in Gelman, 2015b) is connected with the sort of theoretical vagueness that allows nearly unlimited forking paths in data coding, analysis, and interpretation of results.

3 Implications for social science if the model is wrong

Before getting to the implications of the research positions listed above, we will briefly go over some of the methodological problems that led us to discount the claimed evidence.

Feedback loop of credulity. Belief in large effects, naively optimistic power analyses, the building-up of a literature based on apparent "conceptual replications"—this is how bad ideas can stay in circulation. The point here is not that all published work is wrong, or even that all the work above is wrong—a claim can have weak evidential support but still be true—but rather that the normal process of scientific review is not self-correcting, or at least not in the short term.

Researcher degrees of freedom and forking paths. In any given study, there are typically many different choices involved in the coding, inclusion/exclusion, and analysis of data, hence many ways to obtain statistical significance even in the absence of any underlying effect. Simmons, Nelson and Simonsohn (2011) famously labeled this problem as "p-hacking" but we prefer the term "forking paths" to emphasize that the process of adapting an analysis to reach statistical significance can happen without the conscious choice of the researcher, and even if only one analysis is performed on the data at hand (Gelman and Loken 2014).

The statistical significance filter. When only statistically significant results are reported or emphasized, the resulting product systematically leads to overestimation of effect sizes and exaggeration of the importance of claimed effects (e.g., Gerber, Green and Larimer 2008).

Measurement issues. The measurements used in experimental studies are often distant from real-life exposures of interest. Some examples above include questionable claims about dates of ovulation and upper body strength and the questionable relevance of studying a fleeting image on a computer screen. If the effects in question were real and replicable, then the measurement problems would not be such a concern—for example, if subliminal smiley faces

[3]This phrase came from a bestselling book by Kahneman (2011), who later changed his mind, as he explained in this blog comment in 2017: https://replicationindex.com/2017/02/02/reconstruction-of-a-train-wreck-how-priming-research-went-of-the-rails/comment-page-1/#comment-1454

really did have their advertised effects, then this would tell us something about the susceptibility of political opinions to irrelevant stimuli. But given existing statistical concerns, these measurement problems give us one more reason to be skeptical.

Faith in statistics. There seems to be a widespread attitude that causal identification (through experiment or an analysis method in an observational study) plus statistical significance implies near certainty. This belief is mistaken, partly because experiments and identification strategies are imperfect in the real world, and partly because of the forking-path problem listed above.

Social and economic incentives. Direct incentives for researchers to publish, journals to promote dramatic claims and to minimize concerns about the quality of published work, universities to publicize and exaggerate, and journalists to report on breakthroughs. Also selection: more scrupulous researchers are less likely to make big claims; more modest claims are less likely to be published in top journals; more restrained public relations officers are less likely to garner attention for research they are promoted; and more skeptical journalists are less likely to bother with shaky claims. There is also the never-admit-error, no-retraction-no-correction culture.

3.1 Incoherence of the published literature as a whole

Many published claims of large and consistent efffects of small interventions interfere with each other—or they would interfere, if these effects all truly existed. The science of subliminal effects eats itself. If subliminal or even unintentional messages can cause large changes in behavior, then many experimental findings could be explained as byproducts of embodied cognition or social priming. For example, perhaps the instructions given to participants in the Obama-voting study happened to stimulate positive feelings in ovulating women. Similarly, perhaps the results on the subliminal image study had nothing to do with smiley faces but instead involved the triggering of some implicit bias. Or maybe the apparent effects of shark attacks were actually driven by the results of local sporting events.

Typically, when one of these effects is studied, it is without considering all the others, which suggests that researchers when studying effect A, don't really believe in B, C, D, etc.—or else they trust that these other effects will cancel out as a result of the randomization or natural experiment underlying their causal identification. But, conditional on factors B, C, D, etc., having large effects, it is incorrect to assume they can be ignored, even in a randomized study of factor A. The problem is interactions: if the effect of A has large interactions with B, C, etc., then any observed effect is unstable unless the analysis controls or adjusts for all these other factors (Tosh et al. 2025). And this is not merely a theoretical concern, given that such large interactions have been posited based on marital status, parental socioeconomic status, and outdoor air temperature, as well as the topic of study and the location of the experiment. Our guess is that, when interpreting any particular result, researchers do not think so hard about interactions outside of their design, instead putting too much trust in causal identification and statistical significance, without recognizing the assumptions underneath.

3.2 Relation to theories of social and political attitudes and behavior

In the social sciences, the push-a-button, take-a-pill model manifests itself in a belief that people can be given *large and consistent changes* based on small interventions which should be irrelevant, for example, that evacuation behavior is affected by whether a hurricane has a boy or girl name. This, in turn, suggests a model of attitudes and behavior based on indirection, in a way somewhat reminiscent of Freudian analysis. Also in common with the Freudian approach is that the expert—the psychoanalyst, or the professor of psychology—understands people way that may seem contrary to intuition but is consistent with certain deeper views of human nature. In some ways, theories of priming have the worst of both worlds, with the logical looseness that keeps these theories unfalsifiable, but without the psychological insight of Freud. For all the weaknesses of psychoanalytic theory, at least it supposes that the hidden desires which drive our actions are deep and important, unlike social priming theory, which supposes large effects from trivial interventions.

Our problem comes in many of the details and can be summarized by a disquiet with the claimed magnitudes of effects. It is plausible that *some* voters change their political views according to time of the month, but not so plausible that 10% or 20% of women switch in that way. These questions are empirical as well as theoretical, and we do not claim that our skepticism is a strong argument against claims of large effects; rather, given the empirical weaknesses of such claims, it is helpful to remember that they have no compelling theoretical justification.

Ironically, the "behavioral" approach to economics and social science echoes the nearly century-old theory of behaviorism in psychology, a model which was overthrown in the "cognitive revolution" of the 1950s. Considering people as black boxes can seem like pure, objective science—just ask B.F. Skinner—but a huge amount is lost by not directly observing and modeling internal processes. Which can be done in many ways, including asking participants in experiments what they are thinking and involving them more actively in research questions (rather than, as is often done, attempting to use deception to decouple participants' cognitive processes from their decision making).

We can contrast the claims about priming to claims about framing of issues in political debates. As Druckman (2004) discusses, framing effects are considered to be evidence of irrationality in decision making. But issue framing is, if not formally "rational" in the sense of classical utility theory, still *reasonable*, especially when considering attitudes and behaviors in a political context, in which political actors on different sides have incentives to influence people.[4] By contrast, much of the priming literature involves stimuli ("primes") that have little direct relevance to the targeted attitudes and behaviors.

To elaborate on these problems, we consider some of the literature on voter turnout experiments.

[4]We distinguish here between issue framing and "equivalence framing." Framing of the latter type occurs when people's preferences depend on the use of different but logically equivalent phrases. See Chong and Druckman (2007, pp. 114,).

4 Voter turnout

It is now widely accepted that, under a wide range of conditions, interventions designed to increase voter turnout have effects that usually range from zero to four percentage points (Gerber, Huber and Fang 2020). Interventions that raise turnout by more than four percentage points are highly unusual. They tend to depend on striking messages—for example, the "social pressure" messages used in Gerber, Green and Larimer (2008)—or on interventions that are otherwise highly unusual and resource-intensive. Even in these cases, no techniques have been found that consistently raise voter turnout by more than four percentage points. Increasing voter turnout is a difficult business.

Thanks to the rise of large-scale field experiments on voter turnout, most social scientists now agree with this assessment. But not everyone agrees.

Some maintain, instead, that small changes in stimuli—even irrelevant stimuli—can have massive effects on turnout. Two studies in this vein are especially noteworthy: those of Greenwald et al. (1987); Bryan et al. (2011). Consider first the Greenwald et al. (1987) study. The authors find that simply asking people to predict whether they will vote in an upcoming election increases turnout by more than 20 percentage points.

Now consider Bryan et al. (2011). The authors argue that subtle linguistic cues can affect voter turnout. In three experiments, they give subjects surveys that refer to voting with a self-relevant noun ("How important is it to you to be a voter in the upcoming election?") or with a verb ("How important is it to you to vote in the upcoming election?") They find that the noun treatment increases turnout by 11% in California during the 2008 presidential election (Study 2, $N = 88$) and by 14% in New Jersey during the 2009 gubernatorial election (Study 3, $N = 214$). Unlike many studies in this vein, the authors rely on official turnout records to support claims about the effectiveness of the noun treatment. The authors' explanation for the finding is that the noun treatment [activates? enhances?] subjects' tendency to think that being "a voter" is a meaningful part of their identity, which in turn makes them more likely to vote.

These findings were sharply at odds with the turnout literature discussed above. Indeed, they imply that the noun treatment may be the most powerful turnout intervention ever devised. Apparently skeptical of this result, Gerber et al. (2016) used the noun treatment in a large field experiment and found that it had approximately no effect. Some of the authors of the original article wrote a response that the Gerber et al. replication attempt was "inauthentic" because the elections that Gerber et al. had studied were not competitive. In uncompetitive elections, the authors reasoned, voter identity isn't important. In turn, the noun treatment should not be expected to increase turnout in such elections.

In response, Gerber, Huber and Fang (2018) conducted a new field experiment in which the noun treatment was used in more competitive races. They again found approximately no effect. Bryan, Yeager and O'Brien (2019) responded that Gerber, Huber, and Fang had mis-analyzed and misrepresented their own experiment. Specifically, their "more competitive" races were still not all competitive, and their finding of approximately no effect was due to the use of an unusual set of control variables. Use almost any other set of control variables, Bryan et al. claimed, and one will estimate large effects of the noun treatment.

Gerber, Huber, and Fang did not respond directly to all of the charges leveled against their work by Bryan, Yeager and O'Brien (2019). Instead, they fielded yet another field experiment—this one preregistered and fielded in still-more-competitive races, thereby speaking to some of the questions. They again found no effect (Gerber, Huber and Fang 2020).

5 Discussion

A lot of social scientists seem to turn off their critical thinking when a randomized experiment is involved. They forget about questions of validity (are you measuring something of interest) and reliability (are your measurements stable enough to allow you to learn anything generalizable from your sample). These are general concerns; in social science these mistakes in research methods have let in a tide of claims purporting to support a view of voters being blindly led by their genes and hormones and being predictably influenced by irrelevant stimuli, which in turn lends apparent scientific support to cynical views of politics. Some political cynicism is warranted, but we would argue it applies through pathways more generally accepted to be part of political science, involving ideology, partisanship, interest groups, and distributive politics.

The replication crisis in social science goes far beyond the problems with behaviorist models discussed above. For example, Vivalt (2020) finds that published estimates of the effect of policy interventions are wildly optimistic, a finding which should be no surprise given what we know about selection on statistical significance (see e.g., Button et al. 2013) but which has potentially large implications, both for policy and for future practice of policy evaluation. We would care about this even in the absence of any link to particular theories of attitudes and behavior. But as social scientists we are especially concerned about the implications of certain behaviorist theories that have been so strongly pushed by much of the science establishment and the news media, while not being well supported by data. Recognition of the replication crisis should change not only our views of particular studies but also our larger perspective on citizens' views and political deliberation.

References

Achen, Christopher H and Larry M Bartels. 2002. Blind retrospection: Electoral responses to droughts, flu, and shark attacks. In *Annual Meeting of the American Political Science Association*.

Ashworth, Scott and Ethan Bueno De Mesquita. 2014. "Is voter competence good for voters?: Information, rationality, and democratic performance." *American Political Science Review* 108(3):565–587.

Bartels, Larry. 2014. "Here's how a cartoon smiley face punched a big hole in democratic theory.".

Bishop, George F. 2004. *The Illusion of Public Ppinion: Fact and Artifact in American Public Opinion Polls*. Rowman & Littlefield Publishers.

Bryan, Christopher J, David S Yeager and Joseph M O'Brien. 2019. "Replicator degrees of freedom allow publication of misleading failures to replicate." *Proceedings of the National Academy of Sciences* 116(51):25535–25545.

Bryan, Christopher J, Gregory M Walton, Todd Rogers and Carol S Dweck. 2011. "Motivating voter turnout by invoking the self." *Proceedings of the National Academy of Sciences* 108(31):12653–12656.

Bullock, John G and Gabriel Lenz. 2019. "Partisan bias in surveys." *Annual Review of Political Science* 22:325–342.

Button, Katherine S, John Ioannidis, Claire Mokrysz, Brian A Nosek, Jonathan Flint, Emma SJ Robinson and Marcus R Munafò. 2013. "Power failure: Why small sample size undermines the reliability of neuroscience." *Nature Reviews Neuroscience* 14(5):365–376.

Caplan, Bryan. 2011. *The myth of the Rational Voter*. Princeton University Press.

Carter, Travis J, Melissa J Ferguson and Ran R Hassin. 2011. "A single exposure to the American flag shifts support toward Republicanism up to 8 months later." *Psychological Science* 22(8):1011–1018.

Chong, Dennis and James N Druckman. 2007. "Framing theory." *Annual Review of Political Science* 10:103–126.

Druckman, James N. 2004. "Political preference formation: Competition, deliberation, and the (ir) relevance of framing effects." *American political science review* 98(4):671–686.

Durante, Kristina M, Ashley Rae and Vladas Griskevicius. 2013. "The fluctuating female vote: Politics, religion, and the ovulatory cycle." *Psychological Science* 24(6):1007–1016.

Erisen, Cengiz, Milton Lodge and Charles S Taber. 2014. "Affective contagion in effortful political thinking." *Political Psychology* 35(2):187–206.

Ferguson, Thomas. 1995. *Golden Rule: The Investment Theory of Party Competition and the Logic of Money-Driven Political Systems*. University of Chicago Press.

Fowler, Anthony and Andrew B Hall. 2018. "Do shark attacks influence presidential elections? Reassessing a prominent finding on voter competence." *Journal of Politics* 80(4):1423–1437.

Gelman, Andrew. 2015*a*. "Disagreements about the strength of evidence." *Chance* 1:55–59.

Gelman, Andrew. 2015*b*. "Statistics and the crisis of scientific replication." *Significance* 12(3):23–25.

Gelman, Andrew and Eric Loken. 2014. "The statistical crisis in science: data-dependent analysis—a 'garden of forking paths'—explains why many statistically significant comparisons don't hold up." *American Scientist* 102(6):460.

Gerber, Alan, Gregory Huber and Albert Fang. 2018. "Do subtle linguistic interventions priming a social identity as a voter have outsized effects on voter turnout? Evidence from a new replication experiment." *Political Psychology* 39(4):925–938.

Gerber, Alan S, Donald P Green and Christopher W Larimer. 2008. "Social pressure and voter turnout: Evidence from a large-scale field experiment." *American Political Science Review* 102(1):33–48.

Gerber, Alan S, Gregory A Huber and Albert H Fang. 2020. "Voting behavior is unaffected by subtle linguistic cues: Evidence from a psychologically authentic replication." *Behavioural Public Policy* pp. 1–15.

Gerber, Alan S, Gregory A Huber, Daniel R Biggers and David J Hendry. 2016. "A field experiment shows that subtle linguistic cues might not affect voter behavior." *Proceedings of the National Academy of Sciences* 113(26):7112–7117.

Greenwald, Anthony G. 1968. Cognitive learning, cognitive response to persuasion, and attitude change. In *Psychological Foundations of Attitudes*, ed. Anthony G. Greenwald, Timothy C. Brock and Thomas M. Ostrom. New York: Academic Press pp. 147–170.

Greenwald, Anthony G, Catherine G Carnot, Rebecca Beach and Barbara Young. 1987. "Increasing voting behavior by asking people if they expect to vote." *Journal of Applied Psychology* 72(2):315.

Hirano, Shigeo, James M Snyder Jr, Stephen Daniel Ansolabehere and John Mark Hansen. 2010. "Primary elections and partisan polarization in the US Congress." *Quarterly Journal of Political Science* .

Kahneman, Daniel. 2011. *Thinking, Fast and Slow*. Macmillan.

Kahneman, Daniel and Amos Tversky. 1982. "The psychology of preferences." *Scientific American* 246(1):160–173.

Kalla, Joshua L and David E Broockman. 2020. "Reducing exclusionary attitudes through interpersonal conversation: Evidence from three field experiments." *American Political Science Review* 114(2):410–425.

Lax, Jeffrey R and Justin H Phillips. 2009. "Gay rights in the states: Public opinion and policy responsiveness." *American Political Science Review* 103(3):367–386.

Petersen, Michael Bang, Daniel Sznycer, Aaron Sell, Leda Cosmides and John Tooby. 2013. "The ancestral logic of politics: Upper-body strength regulates men's assertion of self-interest over economic redistribution." *Psychological science* 24(7):1098–1103.

Sands, Melissa L. 2017. "Exposure to inequality affects support for redistribution." *Proceedings of the National Academy of Sciences* 114(4):663–668.

Simmons, Joseph P, Leif D Nelson and Uri Simonsohn. 2011. "False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant." *Psychological Science* 22(11):1359–1366.

Szászi, Barnabás, Anthony C. Higney, Aaron B. Charlton, Andrew Gelman, Ignazio Ziano, Balacs Aczel, Daniel G. Goldstein, David S. Yeager and Elizabeth Tipton. 2022. "No reason to expect large and consistent effects of nudge interventions." *Proceedings of the National Academy of Sciences* 119:e2200732119.

Thaler, Richard H and Cass R Sunstein. 2008. *Nudge: Improving Decisions About Health, Wealth, and Happiness*. Yale University Press.

Tosh, Christopher, Philip Greengard, Ben Goodrich, Andrew Gelman, Aki Vehtari and Daniel Hsu. 2025. "The piranha problem: Large effects swimming in a small pond." *Notices of the American Mathematical Society* 72:15–25.

Uhlaner, Carole J. 1989. "Rational turnout: The neglected role of groups." *American Journal of Political Science* pp. 390–422.

Vivalt, Eva. 2020. "How much can we generalize from impact evaluations?" *Journal of the European Economic Association* 18(6):3045–3089.

Zaller, John R. 1992. *The Nature and Origins of Mass Opinion*. Cambridge University Press.