

A Note on Dropping Experimental Subjects who Fail a Manipulation Check

Peter M. Aronow, Jonathon Baron and Lauren Pinson

Yale University, Political Science and Biostatistics, 77 Prospect Street, New Haven, Connecticut, 06520, USA.
Email: peter.aronow@yale.edu

Abstract

Dropping subjects based on the results of a manipulation check following treatment assignment is common practice across the social sciences, presumably to restrict estimates to a subpopulation of subjects who understand the experimental prompt. We show that this practice can lead to serious bias and argue for a focus on what is revealed without discarding subjects. Generalizing results developed in Zhang and Rubin (2003) and Lee (2009) to the case of multiple treatments, we provide sharp bounds for potential outcomes among those who would pass a manipulation check regardless of treatment assignment. These bounds may have large or infinite width, implying that this inferential target is often out of reach. As an application, we replicate Press, Sagan, and Valentino (2013) with a design that does not drop subjects that failed the manipulation check and show that the findings are likely stronger than originally reported. We conclude with suggestions for practice, namely alterations to the experimental design.

Keywords: causal inference, survey experiments, randomized experiments

1 Introduction

Manipulation checks are a valuable means of assessing the robustness of experimental results in studies based on subjects' attention to treatments, for instance, treatment frames presented in survey experiments. In some studies, researchers may be inclined to exclude those subjects who fail the manipulation check from further analysis. This practice is common across the social sciences: we found 59 articles and 36 dissertations that we verified to have dropped subjects after a manipulation (or other posttreatment) check.¹ Articles including this practice have been published in top journals in multiple disciplines over recent years, including the *American Political Science Review*, the *Journal of Personality and Social Psychology*, *Psychological Science*, and the *Journal of Marketing*.² Nominally, the goal of removing subjects is to make sure that we restrict our

Authors' note: Peter M. Aronow is Assistant Professor, Departments of Political Science and Biostatistics, Yale University, 77 Prospect St., New Haven, CT 06520, USA (Email: peter.aronow@yale.edu). Jonathon Baron is Doctoral Student, Department of Political Science, Yale University, 115 Prospect St., New Haven, CT 06511, USA. Lauren Pinson is Doctoral Student, Department of Political Science, Yale University, 115 Prospect St., New Haven, CT 06511, USA. Author names are in alphabetical order and do not reflect relative contributions, which the authors consider to be equal. We thank Allan Dafoe, Don Green, Daniel Masterson, Ben Miller, Molly Offer-Westort, and Betsy Levy Paluck for helpful comments and conversations. Special thanks to Daryl Press, Scott Sagan, and Ben Valentino for generous assistance and materials in replication. We also thank the Yale Institution for Social and Policy Studies Summer Research Lunch group for valuable feedback. Replication data are available in Aronow, Baron, and Pinson (2018).

- The supplementary information provides an extensive bibliography of studies and dissertations that drop or otherwise statistically condition on posttreatment manipulation checks. Articles were found using a series of searches on Google Scholar for "experiment manipulation check," "experiment manipulation attention check," "experiment manipulation attention check political science," "'manipulation check' 'attention check' screen*," "political science 'manipulation check' 'attention check' screen*"; searches for dissertations were performed on ProQuest using "experiment manipulation attention check," which was the most inclusive search on Google Scholar. Articles suspected to use manipulation or posttreatment attention checks as a statistical conditioning strategy were then coded independently by two readers. With the exception of dissertations, when either reader was unsure about how the study was conducted or the readers disagreed, the authors of the article (starting with the corresponding author) were e-mailed for clarification. We sent e-mails to authors regarding 42 articles, all of which received responses, and 28 of which were confirmed to drop subjects based on a manipulation (or other posttreatment) check.
- For recent examples in political science, see Maoz (2006), Small, Lerner, and Fischhoff (2006), Turner (2007), De Oliveira, Guimond, and Dambrun (2012), Crawford *et al.* (2013), and Hoffman *et al.* (2013).

Political Analysis (2019)
vol. 27:572–589
DOI: 10.1017/pan.2019.5

Published
30 May 2019

Corresponding author
Peter M. Aronow

Edited by
J. Grimmer

© The Author(s) 2019. Published
by Cambridge University Press
on behalf of the Society for
Political Methodology.

estimates to a population of subjects who understand the experimental prompt (Wilson, Aronson, and Carlsmith 2010, p. 66) and follow instructions diligently (Oppenheimer, Meyvis, and Davidenko 2009). However, this practice may lead to serious bias in estimation, as dropping subjects may induce an asymmetry across treatment arms.

In this note, we show that the practice of dropping subjects based on a manipulation check should generally be avoided. We provide a number of statistical results establishing that doing so can bias estimates or undermine identification of causal effects. We also show that this practice is equivalent to inducing differential attrition across treatment arms, which may induce bias of unknown sign and magnitude.³ We do not claim that our statistical formulations are particularly novel—they follow from well-known results about conditioning on posttreatment variables and attrition—but, given the prevalence of this practice, we believe that the relationship between these findings and practice in experimentation is underappreciated.

Our contribution is not solely negative—we provide a number of positive results. First, we reiterate the well-known result that the *intent-to-treat* effect is point identified: if subjects are not discarded, a well-defined causal effect can be estimated consistently. Furthermore, we show that when the result of the manipulation check does not depend on the treatment, an alternative causal quantity—the average treatment effect among those who would pass the manipulation check under all conditions—may be estimated dropping subjects. This condition can be ensured in the design of an experiment, by conditioning solely on checks that are delivered before the experimental treatment is administered. When this condition fails, we provide sharp bounds for the average treatment effect among those who would pass the manipulation check under all conditions. Taken together, our results suggest extreme caution in dropping subjects who fail a manipulation check.

In elaborating the potential pitfalls of dropping subjects who fail a manipulation check, we consider Press, Sagan, and Valentino (2013)'s (henceforth PSV) survey experiment on public opinion about nuclear weapons. We provide a number of results from an augmented replication of PSV that does not drop subjects that failed the manipulation check.⁴ Our findings do not contradict the primary substantive findings of PSV, but instead reinforce its claims: we find that study's exclusion of subjects who failed the manipulation check produced weaker findings than would likely have been returned by a full sample. We then conclude with recommendations for applied practice, namely a focus on what is revealed without discarding subjects.

2 Results

Suppose we have an i.i.d. sample from (Y, S, Z) , where Y denotes the subject's response, S denotes the result of a manipulation check (1 if the subject passed, 0 if the subject failed), and Z denotes the subject's treatment assignment (1 for treatment 1, 2 for treatment 2, . . .). Without loss of generality, assume that the support of Z is $\{1, \dots, K\}$, where K is finite.

We make three assumptions to proceed. First, we assume that both potential responses and potential results from the manipulation check are stable, by invoking SUTVA (Rubin 1980), which implies both no interference between units and no multiple unobserved versions of the treatment.

ASSUMPTION 1 (SUTVA). $Y = \sum_{z=1}^K Y(z)\mathbf{I}(Z = z)$ and $S = \sum_{z=1}^K S(z)\mathbf{I}(Z = z)$.

Second, we assume that the treatment is not systematically related to potential outcomes or potential manipulation check results, as would be ensured by random assignment of the treatment.

- 3 The point has been made before, but has not to our knowledge been formalized. For example, Gerber and Green (2012, p. 212) note that attrition may be induced when “[r]esearchers deliberately discard observations. Perhaps ill-advisedly, laboratory researchers sometimes exclude from their analysis subjects who seem not to understand the instructions or who fail to take the experimental situation seriously” but does not provide further discussion of this point.
- 4 Replication data are available from Aronow, Baron, and Pinson (2018).

ASSUMPTION 2 (Ignorability). For all $z \in \{1, \dots, K\}$, $(Y(z), S(z)) \perp\!\!\!\perp Z$, with $\Pr(Z = z) > 0$.

Assumption 2 can be ensured at the design stage by randomizing treatment assignment across subjects.

Finally, we require that at least some subjects pass the manipulation check in both treatment and control.

ASSUMPTION 3 (Nonzero Passing Rates). For all $z \in \{1, \dots, K\}$, $\Pr[S|Z = z] > 0$.

Note that Assumption 3 is verifiable from the experimental data.

2.1 Identification

Without discarding subjects, all mean potential outcomes are well identified, and their differences are also point identified. These differences are sometimes referred to as intent-to-treat effects (Gerber and Green 2012). As we proceed, plug-in estimators will be consistent given suitable regularity conditions (e.g., finite third moments, continuity), with the bootstrap providing a basis for asymptotic inference.

LEMMA 1. $E[Y|Z = z] - E[Y|Z = z'] = E[Y(z) - Y(z')]$.

A proof follows from linearity of expectations. Randomization ensures intent-to-treat effects are point identified, and can be estimated simply by examining differences in means.

In order to assess the operating characteristics of dropping subjects, we must formalize the presumed inferential target of a researcher who chooses to drop subjects based on a manipulation check. Here we consider one possible target that seems natural: $E[Y(z) - Y(z')|S(1) = S(2) = \dots = S(K) = 1]$, or the average treatment effect among subjects who would pass under all treatment conditions. This target parameter is logically equivalent to the complier average causal effect (Angrist, Imbens, and Rubin 1996) when $\text{Supp}(Z) = \{0, 1\}$ and “take-up” is considered to be $D = SZ + (1 - S)(1 - Z)$. However, Angrist, Imbens, and Rubin (1996)’s LATE Theorem, facilitating identification of the target parameter using Z as an instrumental variable for D , is unlikely to hold in this setting. The LATE Theorem requires an exclusion restriction—namely that the effect of the treatment is completely mediated through take-up (e.g., the treatment may have effects on the outcome even for subjects that fail the manipulation check). We thus recommend some caution in instrumental variable-type strategies for utilizing manipulation checks: the validity of the identifying assumptions is not directly verifiable from the experimental data nor can they generally be ensured by design.

There is a condition under which dropping subjects who fail a manipulation check recovers this quantity, namely that in which the treatment assigned to a subject is statistically unrelated to whether or not that subject passes or fails the manipulation.

COROLLARY 1. *If $\Pr[S(1) = S(2) = \dots = S(K)] = 1$, then $E[Y|S = 1, Z = z] - E[Y|S = 1, Z = z'] = E[Y(z) - Y(z')|S(1) = S(2) = \dots = S(K) = 1]$.*

Corollary 1 implies sufficient conditions for discarding subjects to be unproblematic if the inferential target is the average potential outcomes among those who pass the manipulation check. In short, the treatment cannot affect whether or not a subject passes the manipulation check (e.g., if the treatment impacts subjects’ ability to pass the manipulation check itself, for instance by inducing variable degrees of stress, or even if subjects receive treatments for different lengths of time). Thus we can find cases where dropping subjects is acceptable: if a check precedes treatment (e.g., a pretreatment attention check), then discarding subjects is not a problem, at least for characterizing effects for a well-defined subpopulation of units.

Note that the condition in Corollary 1 has a testable implication.

COROLLARY 2. *If $\Pr[S(1) = S(2) = \dots = S(K)] = 1$, then $S \perp\!\!\!\perp Z$.*

A proof follows directly from Assumptions 1 and 2. That is, if the treatment does not affect whether a subject passes the manipulation check, then it must be the case that passing will be unrelated to the treatment condition. Thus, if $S \perp\!\!\!\perp Z$ (i.e., there is an effect of Z on S), then we know that the condition for Corollary 1 to hold (and dropping subjects to be unproblematic) must be false.

However, the converse does not generally hold. In order to justify discarding subjects, it is not sufficient to show that Z and S are unrelated.

COROLLARY 3. $S \perp\!\!\!\perp Z$ does not imply that $E[Y|S = 1, Z = z] - E[Y|S = 1, Z = z'] = E[Y(z) - Y(z')|S(1) = S(2) = \dots = S(K) = 1]$.

Corollary 3 reinforces an important point: even if the failure rates are identical across treatments, conditioning on a manipulation check may still induce bias. This is because the types of subjects who fail the manipulation check under one treatment may not be the same as those who fail under a different treatment.

To this end, potential outcomes among those who would pass regardless of condition are not generally point identified. In a generalization of Lee (2009) (which imposes a monotonicity assumption) and Zhang and Rubin (2003), we derive sharp bounds on potential outcome means $E[Y(z)|S(1) = 1, S(2) = 1, \dots, S(K) = 1]$.⁵

PROPOSITION 1. Suppose that $\Pr[S(1) = S(2) = \dots = S(K) = 1] > 0$ and that Y is continuous with unbounded support. Let $Q_{Y|Z=z, S=1}(\cdot)$ denote the conditional quantile function of Y given $Z = z$ and $S = 1$. Then, sharp bounds for $E[Y(z)|S(1) = S(2) = \dots = S(K) = 1]$ are given by

$$\begin{aligned} E\left[Y|Y \leq Q_{Y|Z=z, S=1}\left(1 - \sum_{z' \in \{1, \dots, K\}: z' \neq z} \frac{\Pr[S = 0|Z = z']}{\Pr[S = 1|Z = z]}\right), Z = z\right] \\ \leq E[Y(z)|S(1) = S(2) = \dots = S(K) = 1] \leq \\ E\left[Y|Y \geq Q_{Y|Z=z, S=1}\left(\sum_{z' \in \{1, \dots, K\}: z' \neq z} \frac{\Pr[S = 0|Z = z']}{\Pr[S = 1|Z = z]}\right), Z = z\right] \end{aligned}$$

when $\sum_{z' \in \{1, \dots, K\}: z' \neq z} \frac{\Pr[S=0|Z=z']}{\Pr[S=1|Z=z]} < 1$, else these bounds are given by $-\infty \leq E[Y(z)|S(1) = S(2) = \dots = S(K) = 1] \leq \infty$.

Assuming that $\Pr[S(1) = S(2) = \dots = S(K) = 1] > 0$ ensures that $E[Y(z)|S(1) = S(2) = \dots = S(K) = 1]$ exists and continuity of Y ensures that the quantile function is well defined. As with Lee (2009)'s bounds, even when Y is discrete, then bounds can be constructed simply by trimming the observations associated with the upper or lower $\sum_{z' \in \{1, \dots, K\}: z' \neq z} \frac{\Pr[S=0|Z=z']}{\Pr[S=1|Z=z]}$ th proportions of the empirical distributions of subjects who pass the manipulation check under treatment and control. With weighted data, this entails using the weighted empirical distribution function. Note that our bounds provide information about the potential for bias introduced by dropping subjects. Differences between the asymptotic value of an estimator computed after dropping subjects and the bounds for $E[Y(z)|S(1) = S(2) = \dots = S(K) = 1]$ provide a range of potential values for asymptotic bias.

2.2 Simulations

In Appendix C, we provide full details and results of a set of simulations to evaluate the properties of the difference-in-means estimator after dropping subjects and of the proposed bounds. We summarize our setup and conclusions here. We assume a treatment Z with $\text{Supp}(Z) = \{1, 2\}$ and $\Pr(Z = 1) = 1/2$. We generated potential outcomes $Y(1) = Y(2) = \lambda[S(2) - S(1)] + N(0, \sigma)$, and vary

5 We thank Ben Miller for helpful discussions regarding the formulation of Proposition 1.

λ , σ , and the joint distribution of $(S(1), S(2))$. In short, we show, all else equal, that bias tends to increase as the average potential outcomes of subjects who would pass the control manipulation check diverge from those who would pass the treatment manipulation check and as failure rates increase. We further show that the width of the bounds increases as failure rates increase, but also as the variability of potential responses increases. Our results shed light on the conditions under which dropping subjects is most problematic, both in terms of the bias introduced and in terms of the fundamental uncertainty about the target parameter.

2.3 Using Covariates for Diagnostics

Suppose that we had a vector of covariates \mathbf{X} associated with each unit, so that we now have an i.i.d. sample from (Y, S, Z, \mathbf{X}) . We can now write an assumption analogous to Assumption 2.

ASSUMPTION 4 (Ignorability with Covariates). For all $z \in \{1, \dots, K\}$, $(Y(z), S(z), \mathbf{X}) \perp\!\!\!\perp Z$, with $\Pr(Z = z) > 0$.

Note that if Z is randomized and \mathbf{X} is measured prior to administration of the treatment, then Assumption 4 holds by construction. However, if \mathbf{X} is measured posttreatment and Z has causal effects on \mathbf{X} , then Assumption 4 may be violated.

These covariates can shed light on the plausibility of the condition for Corollary 1, $\Pr[S(1) = S(2) = \dots = S(K)] = 1$.

COROLLARY 4. If $\Pr[S(1) = S(2) = \dots = S(K)] = 1$, then $\mathbf{X} \perp\!\!\!\perp Z|S$.

A proof directly follows from Assumptions 1 and 4. One practical implication is that the plausibility of the conditions for Corollary 1 can be verified using a balance test. That is, if $\mathbf{X} \not\perp\!\!\!\perp Z|S$, then it must be the case that $\Pr[S(1) = S(2) = \dots = S(K)] \neq 1$.

Again, however, the converse does not hold. In order to justify discarding subjects, it is not sufficient to have balance on the observed covariates.

COROLLARY 5. $\mathbf{X} \perp\!\!\!\perp Z|S$ does not imply that $E[Y|S = 1, Z = z, \mathbf{X} = \mathbf{x}] - E[Y|S = 1, Z = z', \mathbf{X} = \mathbf{x}] = E[Y(z) - Y(z')|S(1) = S(2) = \dots = S(K) = 1, \mathbf{X} = \mathbf{x}]$.

A proof directly follows from that of Corollary 3. So while covariates can help in testing a particular identification condition that justifies dropping subjects, failure to reject the null hypothesis of $\mathbf{X} \perp\!\!\!\perp Z|S$ (or even accepting said null hypothesis) need not imply that dropping subjects is unproblematic.

Note that when Assumption 4 holds, tighter bounds can be obtained than those of Proposition 1. Derivation of these bounds goes outside the scope of the current manuscript, but we note that with even one continuous variable in \mathbf{X} , nonparametric estimation of these bounds will likely be practically difficult, requiring estimation of an infinite-dimensional function.

2.4 Summary of Results

Taken together, our results establish the following: (i) Intent-to-treat effects are point identified. (ii) Potential outcomes among those who would pass a manipulation check under all conditions are not generally point identified. (iii) Sharp bounds for potential outcomes among those who would pass a manipulation check under all conditions may not have finite width. (iv) Showing that equal proportions of subjects failed the manipulation check across all conditions is not sufficient to justify dropping subjects, because the types of subjects that comprise those groups may differ between treatments. (v) If a check precedes treatment (e.g., a pretreatment attention check), or if its result is otherwise unrelated to treatment assignment, dropping subjects who fail such a check does not lead to bias in estimation of outcomes for those who would pass the manipulation check under all treatment conditions. (vi) As passing rates decrease, the risk of bias from dropping subjects increases, and the width of the bounds grows accordingly. (vii) Covariates may provide

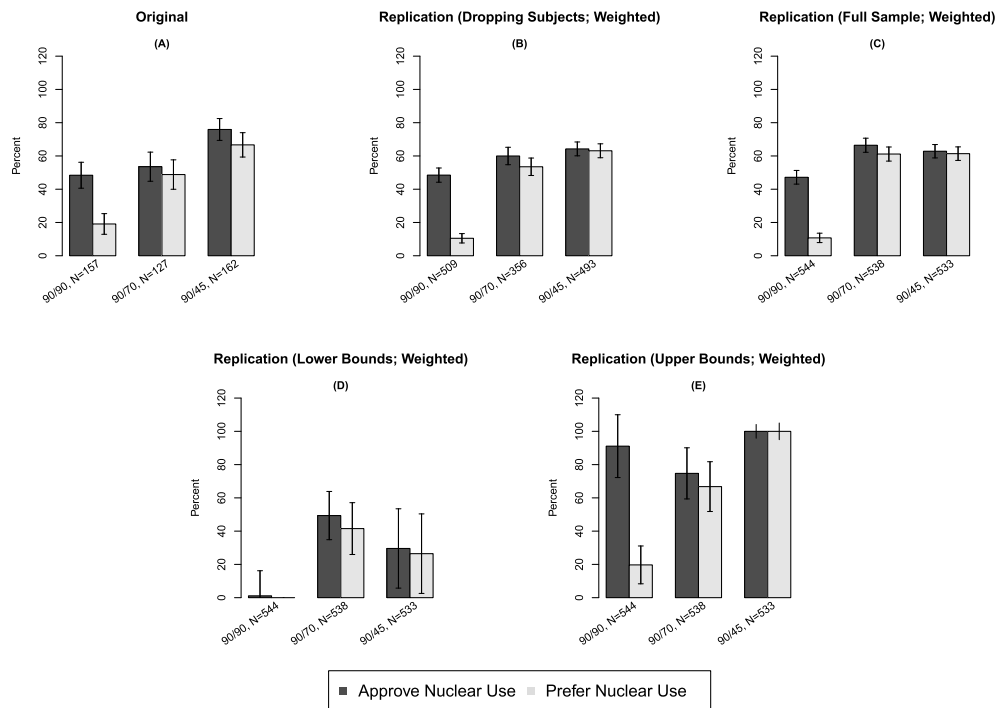


Figure 1. Results from Press, Sagan, and Valentino (2013) and Replication. Comparisons of original and weighted replication data. Panel A presents results from PSV with subjects dropped; Panel B presents results from the replication with subjects dropped; Panel C presents results from the replication using the full sample; Panel D presents results imputing the lower bounds for all treatment conditions; Panel E presents results imputing the upper bounds for all treatment conditions. Vertical bars represent 95% confidence intervals on point estimates calculated using the bootstrap.

information about the consequences of dropping subjects, though their informative power is limited.

3 Application

We now discuss these findings in the context of PSV, which presents a survey conducted using a representative sample of voting-age American citizens. PSV reports on whether public opinion regarding nuclear weapons use is shaped primarily by ethical or strategic considerations. PSV’s five treatment conditions detail a scenario in which an Al Qaeda nuclear weapons lab in Syria obtains weapons-grade uranium intended for offensive use against the United States. We focus on three of the treatments, which manipulate the relative expected effectiveness of a nuclear or conventional strike on the Al Qaeda facility. The treatments describe the effectiveness ratios of nuclear/conventional weapons at 90 percent/90 percent, 90 percent/70 percent, and 90 percent/45 percent, respectively (henceforth referred to as 90–90, 90–70, and 90–45). As demonstrated in Panel A of Figure 1, PSV reports a strongly monotonic increase in subjects’ approval of and preference for nuclear weapons use as nuclear weapons’ effectiveness increases relative to that of conventional weapons.

However, we will show that this finding—namely a strong monotonic increase—is partly attributable to dropping subjects, and that, without dropping subjects, even stronger results are obtained. PSV utilized a manipulation check following the treatment as a means of gauging subjects’ attention to the treatment articles; subjects who fail the manipulation check are dropped from the analysis. This check asked subjects to choose from five options including whether the treatment they had received had concluded that nuclear weapons would be equally effective, moderately more effective, or much more effective than conventional weapons; these answers

were intended to correspond to the 90–90, 90–70, and 90–45 comparisons, respectively. Subjects who failed the manipulation check were given an opportunity to read the article a second time, but their responses were dropped due to failure.

We conducted a successful replication of PSV, showing that the primary findings of the study are robust, including robustness to treatment variations not considered in the original article (details are provided in Appendix B). One notable variation in our replication was that subjects were not informed of whether they had passed or failed the manipulation check, and data were collected regardless, allowing us to assess the consequences of dropping subjects. Our study used 2,733 subjects recruited from Amazon.com Mechanical Turk.⁶ Note that there is a good body of evidence to suggest that subjects on Mechanical Turk tend to be more attentive than other (representative) samples (Hauser and Schwarz 2016), and thus it is possible that our passing rates are higher than those of PSV. If so, then all else equal, the bias in our estimates after dropping subjects would be smaller, and the width of our bounds would also be smaller than the bounds that would be associated with the original data.

Subjects were compensated at \$.50 each, with the added chance of winning a \$100 bonus if they passed the manipulation check. For our primary analysis, we used inverse probability weighting (IPW) to adjust the replication sample to match the covariate distribution of the sample used in PSV, and computed estimates after weighting.⁷ We then performed our analysis both including and excluding subjects who failed the manipulation check.

The primary results of our replication are presented in Panels B and C of Figure 1. (Unweighted analyses are presented in Figure 2.) PSV argues that a clear majority of subjects both approve of and prefer a prospective nuclear strike in only 90–45. However, we show that these findings are actually attenuated as a consequence of dropping the subjects who failed from the analysis. Panel C demonstrates that including data from all subjects, regardless of performance on the manipulation check, alters results substantially, rendering estimates in 90–70 and 90–45 practically indistinguishable. The discrepancy is notable in its substantive importance for the results of PSV: subjects dropped from the analysis actually perceived a “moderate” decline in conventional weapons’ relative effectiveness to be a “significant” decrease; including such subjects in the analysis provides even stronger evidence against a nuclear-nonuse norm than PSV depicts.

The replication also illustrates the importance of Corollary 3: showing the equivalence between failure rates under two conditions does not imply that the types of individuals who fail are equivalent. Treatments 1 and 3 have statistically and substantively indistinguishable failure rates (6.4% vs. 7.5%), and yet the covariate profiles of those who failed the manipulation check are strikingly different, as can be seen in Table 2 in Appendix D. For example, 63.7% of subjects who fail the manipulation check under Treatment 3 are female, but only 24.9% of the subjects who fail the manipulation check under Treatment 1 are female. Similar differences are found in political party, religious importance, region, and racial composition.

We also report sharp bounds for the average potential outcomes among subjects who would pass the manipulation check regardless of treatment assignment. We are unable to calculate analogous bounds using the PSV data, because no information was available on the proportion of subjects who were dropped. Panel D depicts our results imputing the lower bound, whereas Panel E shows our estimates imputing the upper bound. We observe that the bounds for all treatment

6 Three subjects were omitted from analysis because of technical difficulties that prevented us from verifying that they completed the survey; 2,730 subjects are included in the analysis below.

7 Let $R_i = 1$ if an observation i is in the replication study, else let $R_i = 0$ if an observation was in the original data. We performed a logistic regression of R_i on the following covariates \mathbf{X}_i : Education, Party, Religion, Political Interest, Income, Gender, News Interest, Voter Registration, Birth Year, Region, Race, and Ideology (with mean-imputation for missingness). Using the output of this logistic regression, we computed a predicted value $p_i = \Pr[R_i = 1 | \mathbf{X}_i]$ for each observation i . To reweight the replication study to the original study’s covariate profile, we weighted each observation in the replication sample by $\frac{p_i}{1-p_i}$.

conditions have overlapping regions, in both outcome variables, making it impossible to fully differentiate results across the treatments. The bounds are largely uninformative, suggesting the fundamental uncertainty induced by attempting to estimate effects among the subpopulation of subjects who would always pass the manipulation check. If these effects are indeed the inferential target of the researcher, very little is revealed by the experimental data, and dropping subjects may introduce serious bias of unknown sign and magnitude (as any values within the range of the bounds are compatible with the experimental data).

4 Conclusion

We reiterate that our critiques do not apply to research designs that use pretreatment attention checks to screen subjects, as established by Corollary 3. Attention checks placed before treatment can be used to prune subjects from final analysis in a principled manner. Although screening changes the inferential target from the whole population to a subpopulation of subjects who are paying sufficient attention and have the ability to pass the attention check, it does not compromise internal validity. The use of manipulation checks in the pilot stage can provide information for improving the interpretability of treatments and manipulation checks to maximize passing rates in the manipulation. The use of pretreatment checks and piloting may also improve estimates by focusing only on the subpopulation of subjects who are able to pass and do so. Of course, there is significant debate about best practices here (see, e.g., Peer, Vosgerau, and Acquisti 2014). As such, we recommend Oppenheimer, Meyvis, and Davidenko (2009)'s suggestions for compelling experimental subjects to participate more diligently; as well as Berinsky, Margolis, and Sances (2014)'s analysis of the benefits and drawbacks of screening, and recommendations for practice. Oppenheimer, Meyvis, and Davidenko (2009) finds that forcing inattentive subjects to repeat Instructional Manipulation Checks (IMCs) until the subjects read the checks thoroughly homogenizes the subject group. We recognize the potential benefits of this approach when performed prior to treatment, but we caution against the use of posttreatment IMCs as a statistical conditioning strategy. Berinsky, Margolis, and Sances (2014)'s suggested use of screens to assess subjects' attention on a continuum also represents a transparent approach to presenting findings by showing the results conditional on each value of attention. However, doing so with a posttreatment check would again introduce the issues discussed here.

In general, we stress the importance of manipulations that are sufficiently clear so as to minimize the necessity to remove subjects based on a lack of comprehension. Although we have proposed bounds for causal effects among the subjects who would always pass the manipulation check, these bounds will be uninformative in practice when failure rates are high. Pilot studies may help to ensure that the treatments are understood by subjects as the researchers intended, with the caveat that the pilot population may be unrepresentative of the final test population. We underline that best practice should maintain a focus on intent-to-treat effects, which are generally point identified and have a clear substantive interpretation. The credibility of the experiment ultimately rests on the quality of the manipulation, rather than post hoc statistical adjustments. As a means of validation, manipulation checks can help researchers to understand whether or not this criterion has been met. But dropping subjects who fail a manipulation check presented after the intervention may introduce biases of unknown sign and magnitude.

Supplementary material

For supplementary material accompanying this paper, please visit <https://doi.org/10.1017/pan.2019.5>.

Appendix A. Proofs

Proof of Corollary 1. $\Pr[S(1) = S(2) = \dots = S(K)] = 1$ implies $S = S(1) = S(2) = \dots = S(K)$, thus ensuring $(Y(1), Y(2), \dots, Y(K), S) \perp\!\!\!\perp Z$. Joint independence implies that

$$E[Y(z)|S = 1, Z = z] - E[Y(z')|S = 1, Z = z'] = E[Y(z)|S(1) = 1, S(2) = 1, \dots, S(K) = 1] - E[Y(z')|S = 1]. \quad \square$$

Proof of Corollary 3. We prove the claim via a simple counterexample. Suppose $\text{Supp}(Z) = \{1, 2\}$ and

$$f(Y(1), Y(2), S(1), S(2)) = \begin{cases} 1/3 : Y(1) = 1, Y(2) = 1, S(1) = 0, S(2) = 1 \\ 1/3 : Y(1) = 0, Y(2) = 0, S(1) = 1, S(2) = 0 \\ 1/3 : Y(1) = 0, Y(2) = 0, S(1) = 1, S(2) = 1 \\ 0 : \text{otherwise.} \end{cases}$$

Note that $E[S(2) - S(1)] = 0$ and $E[\tau|S(1) = S(2) = 1] = E[\tau|S(1) = 1, S(2) = 0] = E[\tau|S(1) = 0, S(2) = 1] = 0$. $E[Y|S = 1, Z = 2] - E[Y|S = 1, Z = 1] = 0 - 1/2 = -1/2$. \square

Proof of Proposition 1. We will follow the general logic of Lee (2009), and technical details carry through from the proof of Lee's Proposition 1a. Without loss of generality, we consider the upper bound for $E[Y(1)|S(1) = S(2) = \dots = S(K) = 1]$.

Define $U = 1$ if $S(2) = \dots = S(K) = 1$, else let $U = 0$. Then $E[Y(1)|S(1) = S(2) = \dots = S(K) = 1] = E[Y(1)|U = 1, S(1) = 1]$. We do not observe the joint distribution of $(Y(1), U)|S(1) = 1$, as we never jointly observe $Y(z)$ and $S(z')$, for $z \neq z'$. Let $p_{U^0} = \Pr[U = 0|S(1) = 1]$. Given continuity of $Y(1)$, then among all possible joint distributions $(Y(1), U)|S(1) = 1$, $E[Y(1)|U = 1, S(1) = 1]$ is maximized when $U = 1$ for all $Y(1) \geq Q_{Y(1)}(p_{U^0})$. By weak monotonicity of the quantile function, it suffices to maximize p_{U^0} to find a maximum for $E[Y(1)|U = 1, S(1) = 1]$.

We again do not observe the joint distribution $(U, S(1))$. By σ -additivity, a sharp upper bound is obtained for $\Pr[U = 0]$ is obtained when the regions where $S(2), S(3), \dots, S(K)$ each equal zero are disjoint, with

$$\Pr[U = 0] \leq \begin{cases} \sum_{k=2}^K \Pr[S(k) = 0] : \sum_{k=2}^K \Pr[S(k) = 0] < 1 \\ 1 : \sum_{k=2}^K \Pr[S(k) = 0] \geq 1. \end{cases}$$

Thus, among all possible joint distributions $(U, S(1))$, $\Pr[U = 0|S(1) = 1] = p_{U^0}$ is maximized when

$$p_{U^0} = \begin{cases} \frac{\sum_{k=2}^K \Pr[S(k) = 0]}{\Pr[S(1) = 1]} : \frac{\sum_{k=2}^K \Pr[S(k) = 0]}{\Pr[S(1) = 1]} < 1 \\ 1 : \frac{\sum_{k=2}^K \Pr[S(k) = 0]}{\Pr[S(1) = 1]} \geq 1. \end{cases}$$

Thus if $\frac{\sum_{k=2}^K \Pr[S(k)=0]}{\Pr[S(1)=1]} < 1$, a sharp upper bound is given by $E[Y(1)|U = 1, S(1) = 1] \leq E[Y(1)|Y(1) \leq Q_{Y(1)|S(1)=1}(1 - \sum_{k=2}^K \frac{\Pr[S(k)=0]}{\Pr[S(1)=1]})]$, else the upper bound is infinite.

By random assignment and SUTVA, the conditional distribution of $Y(1)|S(1) = 1$ is equivalent to the conditional distribution of $Y|S = 1, Z = 1$, and the marginal distributions of $S(k)$ are each equivalent to $S|Z = k$. Thus a sharp upper bound is given by $E[Y(1)|U = 1, S(1) = 1] \leq E[Y|Y \geq Q_{Y|Z=1, S=1}(\sum_{k=2}^K \frac{\Pr[S=0|Z=k]}{\Pr[S=1|Z=k]})]$, $Z = 1$ when $\sum_{k=2}^K \frac{\Pr[S=0|Z=k]}{\Pr[S=1|Z=k]} < 1$, else the upper bound is infinite. The bounds are invariant to indexing of treatments Z , thus yielding the general upper bound in Proposition 1. Analogous calculations yield lower bounds. \square

Appendix B. Details of Replication of Press, Sagan, and Valentino (2013)

Our replication and preanalysis plan are hosted at EGAP (ID: 20150131AA). Our replication included three major variations, the analysis of which underlines the robustness of PSV. We list these analyses in turn below.

First, because the original experiment was performed prior to the onset of the Syrian civil war, we sought to assess whether the results were invariant to shifts in time and context (i.e., whether the results might differ in our replication, given the political changes that have occurred in Syria). We thus randomized whether treatment frames presented the scenario in Syria or Lebanon, which was used as an analog to pre-civil-war Syria; treatments were assigned through a 2×5 factorial design. We found no statistically or substantively significant difference between Syria and Lebanon treatment frames, demonstrating that the results presented in PSV are robust to these temporal and contextual changes.

Second, we analyzed whether the PSV study's use of posttreatment covariates introduced bias. We added another treatment (rendering our augmented replication a $2 \times 2 \times 5$ factorial design) that randomized whether subjects answered these questions before or after treatment. This analysis failed to reveal any statistically or substantively significant results.

Third, as noted above, we performed weighting on our survey sample to approximate the experimental population used by PSV. Our subjects were recruited from Mechanical Turk, and likely constituted an unrepresentative sample. As noted in the main text, we used logistic regression and IPW to assign treatment probabilities and corresponding weights for each subject. We did observe differences between the weighted and unweighted analyses, but neither undermined the substantive findings of PSV.

Appendix C. Simulations

We assume a treatment Z with $\text{Supp}(Z) = \{1, 2\}$ and $\Pr(Z = 1) = 1/2$. We generated potential outcomes $Y(1) = Y(2) = \lambda[S(2) - S(1)] + N(0, \sigma)$, and vary λ , σ , and the joint distribution of $(S(1), S(2))$. Note that in the simulation, we have assumed that there is no effect of the treatment whatsoever, and the results would be invariant to the introduction of any constant treatment effect. λ represents the divergence in potential outcomes between those who would pass and those who would fail the manipulation check and σ represents the unexplained variability of potential outcomes. To put our results in asymptopia, we assume $N = 1,000$, and perform 100,000 simulations.

Table 1 presents the results of our simulations. We first discuss the bias of the difference-in-means estimator after dropping subjects. We show that bias tends to increase as λ —the divergence between the average potential outcomes of subjects who would pass the control manipulation check and that of those who would pass the treatment manipulation check—increases. See, e.g., row 1 vs. 2. As failure rates increase, not necessarily differentially across treatment arms, we also see that bias increases; compare rows 1–4 to 5–8 to 9–12. Furthermore, as $\rho(S(1), S(2))$ —the correlation between potential responses to the manipulation check—decreases, bias also increases, as evidenced by, e.g., row 4 vs. row 1.

The width of the bounds also depends on multiple factors. As the variability of potential outcomes increases (characterized by σ , and to a lesser extent λ), the width of the bounds increases, as evidenced by comparing, e.g., row 1 vs. 2 vs. 3. The width of the bounds also depends on failure rates; again compare rows 1–4 to 5–8 to 9–12. The bounds do not depend on any unobservable features of the joint distributions of potential outcomes and responses to the manipulation check. To wit, the width of the bounds does not change as $\rho(S(1), S(2))$ is varied; compare, e.g., row 1 to row 4.

Table 1. Simulations demonstrating the effects of dropping. Simulations performed with $N = 1,000$ and 100,000 simulations; bound widths are presented as averages over all simulations.

	λ	$\Pr[S(1) = 1]$	$\Pr[S(2) = 1]$	$\rho(S(1), S(2))$	σ	Bias	Bound Width
1	1	0.8	0.8	0.25	1	0.300	1.753
2	10	0.8	0.8	0.25	1	3.005	5.070
3	1	0.8	0.8	0.25	10	0.298	16.543
4	1	0.8	0.8	0.50	1	0.200	1.723
5	1	0.6	0.8	0.25	1	0.456	1.613
6	10	0.6	0.8	0.25	1	4.561	3.403
7	1	0.6	0.8	0.25	10	0.459	15.247
8	1	0.6	0.8	0.50	1	0.315	1.569
9	1	0.6	0.6	0.25	1	0.598	4.692
10	10	0.6	0.6	0.25	1	5.973	19.712
11	1	0.6	0.6	0.25	10	0.598	42.635
12	1	0.6	0.6	0.50	1	0.398	4.589

Appendix D. Additional (Weighted) Summary Statistics

Below, we present distributions of the reweighted covariate profiles of subjects in our replication study, disaggregated by treatment condition and performance on the manipulation checks.

Table 2. Weighted covariate distributions among subjects who failed the manipulation check.

	Treatment 1		Treatment 2		Treatment 3	
	Mean	(SE)	Mean	(SE)	Mean	(SE)
<i>Education</i>	2.931	0.178	3.236	0.213	2.801	0.168
<i>Democrat</i>	0.336	0.135	0.545	0.086	0.390	0.114
<i>Republican</i>	0.424	0.186	0.355	0.082	0.237	0.088
<i>Independent</i>	0.172	0.103	0.072	0.027	0.168	0.092
<i>Other party</i>	0.042	0.031	0.015	0.008	0.032	0.023
<i>Political interest (5-point scale)</i>	3.542	0.195	3.782	0.114	3.090	0.277
<i>Income</i>	7.558	0.692	6.253	0.491	7.192	0.616
<i>Age (continuous)</i>	40.960	5.781	52.658	3.535	39.028	3.592
<i>Age (1st quartile)</i>	0.470	0.170	0.231	0.049	0.611	0.126
<i>Age (2nd quartile)</i>	0.242	0.127	0.171	0.044	0.107	0.080
<i>Age (3rd quartile)</i>	0.000	0.000	0.214	0.063	0.180	0.107
<i>Age (4th quartile)</i>	0.287	0.203	0.384	0.104	0.102	0.090
<i>Female</i>	0.249	0.114	0.688	0.069	0.637	0.102
<i>Ideology (5-point scale)</i>	3.116	0.318	3.019	0.145	2.874	0.202
<i>News interest (4-point scale)</i>	1.959	0.290	1.636	0.113	2.458	0.265
<i>Registered to vote</i>	0.746	0.119	0.919	0.041	0.832	0.092
<i>White</i>	0.630	0.141	0.867	0.035	0.733	0.095
<i>Black</i>	0.292	0.125	0.053	0.020	0.036	0.027
<i>Other race</i>	0.078	0.048	0.080	0.025	0.231	0.093
<i>Religious importance</i>	4.680	0.457	4.623	0.257	3.774	0.391
<i>Northeast</i>	0.034	0.023	0.121	0.037	0.217	0.104
<i>Midwest</i>	0.056	0.035	0.404	0.099	0.183	0.093
<i>South</i>	0.406	0.149	0.326	0.074	0.511	0.111
<i>West</i>	0.504	0.166	0.149	0.041	0.089	0.042
<i>Elite</i>	0.089	0.073	0.076	0.046	0.059	0.037
<i>Observations</i>	$N = 35$		$N = 182$		$N = 40$	

Table 3. Weighted covariate distributions among subjects who passed the manipulation check.

	Treatment 1		Treatment 2		Treatment 3	
	Mean	(SE)	Mean	(SE)	Mean	(SE)
<i>Education</i>	3.620	0.123	3.579	0.138	3.583	0.127
<i>Democrat</i>	0.538	0.045	0.415	0.053	0.466	0.052
<i>Republican</i>	0.275	0.039	0.436	0.058	0.371	0.055
<i>Independent</i>	0.137	0.027	0.095	0.030	0.122	0.026
<i>Other party</i>	0.041	0.023	0.045	0.021	0.023	0.007
<i>Political interest (5-point scale)</i>	3.649	0.096	3.785	0.103	3.764	0.092
<i>Income</i>	7.867	0.261	7.608	0.403	8.064	0.315
<i>Age (continuous)</i>	47.926	1.328	50.823	1.639	49.849	1.617
<i>Age (1st quartile)</i>	0.310	0.032	0.245	0.032	0.267	0.032
<i>Age (2nd quartile)</i>	0.163	0.028	0.141	0.028	0.162	0.025
<i>Age (3rd quartile)</i>	0.240	0.044	0.266	0.049	0.249	0.044
<i>Age (4th quartile)</i>	0.287	0.050	0.348	0.063	0.321	0.062
<i>Female</i>	0.583	0.043	0.656	0.048	0.594	0.046
<i>Ideology (5-point scale)</i>	2.853	0.101	3.154	0.106	3.091	0.127
<i>News interest (4-point scale)</i>	1.798	0.072	1.788	0.071	1.724	0.078
<i>Registered to vote</i>	0.927	0.024	0.906	0.026	0.927	0.020
<i>White</i>	0.761	0.035	0.817	0.042	0.768	0.045
<i>Black</i>	0.076	0.017	0.079	0.037	0.125	0.039
<i>Other race</i>	0.163	0.031	0.104	0.028	0.107	0.030
<i>Religious importance</i>	3.707	0.161	3.837	0.193	4.285	0.177
<i>Northeast</i>	0.168	0.034	0.196	0.042	0.121	0.023
<i>Midwest</i>	0.230	0.041	0.232	0.052	0.215	0.040
<i>South</i>	0.322	0.040	0.302	0.052	0.493	0.053
<i>West</i>	0.280	0.042	0.270	0.047	0.171	0.030
<i>Elite</i>	0.086	0.030	0.055	0.014	0.091	0.028
<i>Observations</i>	<i>N</i> = 509		<i>N</i> = 356		<i>N</i> = 493	

Table 4. Weighted covariate distributions for all subjects.

	Treatment 1		Treatment 2		Treatment 3	
	Mean	(SE)	Mean	(SE)	Mean	(SE)
<i>Education</i>	3.576	0.113	3.450	0.124	3.535	0.119
<i>Democrat</i>	0.525	0.044	0.464	0.049	0.461	0.050
<i>Republican</i>	0.285	0.040	0.406	0.048	0.363	0.051
<i>Independent</i>	0.139	0.027	0.086	0.021	0.125	0.025
<i>Other party</i>	0.041	0.022	0.034	0.014	0.023	0.007
<i>Political interest (5-point scale)</i>	3.642	0.090	3.784	0.078	3.723	0.090
<i>Income</i>	7.847	0.245	7.098	0.324	8.010	0.298
<i>Age (continuous)</i>	47.480	1.295	51.514	1.725	49.184	1.576
<i>Age (1st quartile)</i>	0.320	0.031	0.239	0.027	0.288	0.032
<i>Age (2nd quartile)</i>	0.168	0.027	0.153	0.024	0.159	0.024
<i>Age (3rd quartile)</i>	0.224	0.041	0.247	0.040	0.245	0.043
<i>Age (4th quartile)</i>	0.287	0.050	0.361	0.056	0.307	0.059
<i>Female</i>	0.562	0.043	0.668	0.039	0.597	0.043
<i>Ideology (5-point scale)</i>	2.870	0.097	3.103	0.088	3.077	0.122
<i>News interest (4-point scale)</i>	1.808	0.070	1.731	0.065	1.769	0.078
<i>Registered to vote</i>	0.916	0.024	0.911	0.023	0.921	0.019
<i>White</i>	0.753	0.034	0.836	0.030	0.766	0.043
<i>Black</i>	0.089	0.018	0.069	0.025	0.119	0.036
<i>Other race</i>	0.158	0.030	0.095	0.020	0.114	0.029
<i>Religious importance</i>	3.769	0.156	4.133	0.167	4.254	0.168
<i>Northeast</i>	0.159	0.033	0.168	0.030	0.127	0.023
<i>Midwest</i>	0.219	0.039	0.297	0.053	0.213	0.039
<i>South</i>	0.328	0.038	0.311	0.043	0.494	0.050
<i>West</i>	0.294	0.042	0.224	0.034	0.166	0.028
<i>Elite</i>	0.086	0.028	0.063	0.019	0.089	0.027
<i>Observations</i>	<i>N</i> = 544		<i>N</i> = 538		<i>N</i> = 533	

Appendix E. Additional (Unweighted) Summary Statistics

Below, we present distributions of the unweighted covariate profiles of subjects in our replication study, disaggregated by treatment condition and performance on the manipulation checks.

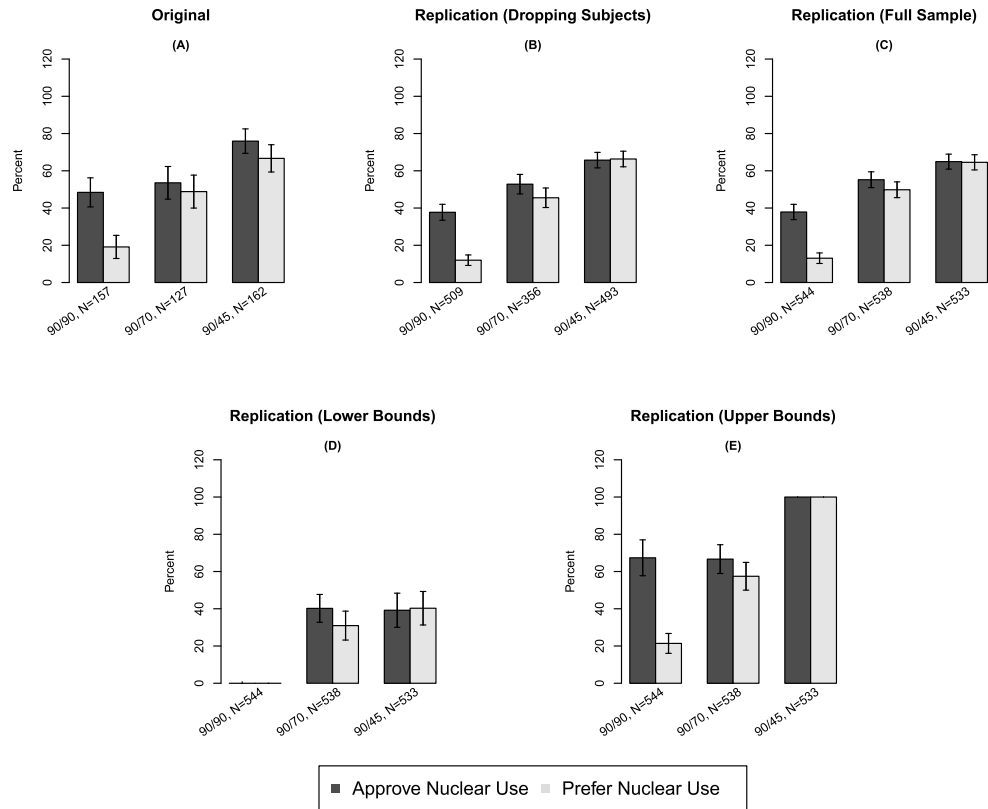


Figure 2. Unweighted Results from Press, Sagan, and Valentino (2013) and Replication. Comparisons of original and unweighted replication data. Panel A presents results from PSV with subjects dropped; Panel B presents results from the replication with subjects dropped; Panel C presents results from the replication using the full sample; Panel D presents results imputing the lower bounds for all treatment conditions; Panel E presents results imputing the upper bounds for all treatment conditions. Vertical bars represent 95% confidence intervals on point estimates calculated using the bootstrap.

Table 5. Unweighted covariate distributions among subjects who failed the manipulation check.

	Treatment 1		Treatment 2		Treatment 3	
	Mean	(SE)	Mean	(SE)	Mean	(SE)
<i>Education</i>	2.931	0.178	3.236	0.213	2.801	0.168
<i>Democrat</i>	0.336	0.135	0.545	0.086	0.390	0.114
<i>Republican</i>	0.424	0.186	0.355	0.082	0.237	0.088
<i>Independent</i>	0.172	0.103	0.072	0.027	0.168	0.092
<i>Other party</i>	0.042	0.031	0.015	0.008	0.032	0.023
<i>Political interest (5-point scale)</i>	3.542	0.195	3.782	0.114	3.090	0.277
<i>Income</i>	7.558	0.692	6.253	0.491	7.192	0.616
<i>Age (continuous)</i>	40.960	5.781	52.658	3.535	39.028	3.592
<i>Age (1st quartile)</i>	0.470	0.170	0.231	0.049	0.611	0.126
<i>Age (2nd quartile)</i>	0.242	0.127	0.171	0.044	0.107	0.080
<i>Age (3rd quartile)</i>	0.000	0.000	0.214	0.063	0.180	0.107
<i>Age (4th quartile)</i>	0.287	0.203	0.384	0.104	0.102	0.090
<i>Female</i>	0.249	0.114	0.688	0.069	0.637	0.102
<i>Ideology (5-point scale)</i>	3.116	0.318	3.019	0.145	2.874	0.202
<i>News interest (4-point scale)</i>	1.959	0.290	1.636	0.113	2.458	0.265
<i>Registered to vote</i>	0.746	0.119	0.919	0.041	0.832	0.092
<i>White</i>	0.630	0.141	0.867	0.035	0.733	0.095
<i>Black</i>	0.292	0.125	0.053	0.020	0.036	0.027
<i>Other race</i>	0.078	0.048	0.080	0.025	0.231	0.093
<i>Religious importance</i>	4.680	0.457	4.623	0.257	3.774	0.391
<i>Northeast</i>	0.034	0.023	0.121	0.037	0.217	0.104
<i>Midwest</i>	0.056	0.035	0.404	0.099	0.183	0.093
<i>South</i>	0.406	0.149	0.326	0.074	0.511	0.111
<i>West</i>	0.504	0.166	0.149	0.041	0.089	0.042
<i>Elite</i>	0.089	0.073	0.076	0.046	0.059	0.037
<i>Observations</i>	<i>N</i> = 35		<i>N</i> = 182		<i>N</i> = 40	

Table 6. Unweighted covariate distributions among subjects who passed the manipulation check.

	Treatment 1		Treatment 2		Treatment 3	
	Mean	(SE)	Mean	(SE)	Mean	(SE)
<i>Education</i>	4.175	0.055	4.099	0.066	4.201	0.055
<i>Democrat</i>	0.568	0.022	0.553	0.026	0.535	0.022
<i>Republican</i>	0.218	0.018	0.233	0.022	0.219	0.019
<i>Independent</i>	0.147	0.016	0.129	0.018	0.156	0.016
<i>Other party</i>	0.043	0.009	0.059	0.012	0.047	0.009
<i>Political interest (5-point scale)</i>	3.285	0.044	3.427	0.049	3.302	0.041
<i>Income</i>	7.214	0.147	7.607	0.177	7.659	0.156
<i>Age (continuous)</i>	33.900	0.530	34.963	0.680	34.572	0.523
<i>Age (1st quartile)</i>	0.739	0.020	0.688	0.025	0.708	0.020
<i>Age (2nd quartile)</i>	0.126	0.015	0.140	0.018	0.162	0.017
<i>Age (3rd quartile)</i>	0.079	0.012	0.101	0.016	0.079	0.012
<i>Age (4th quartile)</i>	0.057	0.010	0.070	0.013	0.051	0.010
<i>Female</i>	0.481	0.022	0.466	0.026	0.472	0.023
<i>Ideology (5-point scale)</i>	2.530	0.046	2.641	0.057	2.631	0.045
<i>News interest (4-point scale)</i>	2.114	0.040	2.037	0.046	2.150	0.042
<i>Registered to vote</i>	0.871	0.015	0.864	0.018	0.876	0.015
<i>White</i>	0.752	0.019	0.815	0.021	0.793	0.018
<i>Black</i>	0.077	0.012	0.045	0.011	0.073	0.012
<i>Other race</i>	0.172	0.017	0.140	0.018	0.134	0.015
<i>Religious importance</i>	2.680	0.084	2.640	0.099	2.828	0.086
<i>Northeast</i>	0.185	0.017	0.219	0.022	0.219	0.018
<i>Midwest</i>	0.238	0.019	0.222	0.022	0.219	0.019
<i>South</i>	0.340	0.021	0.320	0.025	0.337	0.021
<i>West</i>	0.238	0.019	0.239	0.023	0.225	0.019
<i>Elite</i>	0.081	0.012	0.110	0.017	0.099	0.014
<i>Observations</i>	<i>N</i> = 509		<i>N</i> = 356		<i>N</i> = 493	

Table 7. Unweighted covariate distributions for all subjects.

	Treatment 1		Treatment 2		Treatment 3	
	Mean	(SE)	Mean	(SE)	Mean	(SE)
<i>Education</i>	4.129	0.054	4.076	0.054	4.131	0.053
<i>Democrat</i>	0.562	0.022	0.552	0.022	0.522	0.021
<i>Republican</i>	0.215	0.018	0.238	0.018	0.218	0.018
<i>Independent</i>	0.151	0.015	0.132	0.014	0.158	0.016
<i>Other party</i>	0.046	0.009	0.052	0.010	0.049	0.009
<i>Political interest (5-point scale)</i>	3.279	0.041	3.388	0.041	3.288	0.040
<i>Income</i>	7.216	0.142	7.210	0.148	7.595	0.148
<i>Age (continuous)</i>	33.557	0.509	35.277	0.550	34.285	0.507
<i>Age (1st quartile)</i>	0.746	0.019	0.682	0.020	0.719	0.019
<i>Age (2nd quartile)</i>	0.125	0.014	0.147	0.015	0.154	0.016
<i>Age (3rd quartile)</i>	0.074	0.011	0.100	0.013	0.079	0.012
<i>Age (4th quartile)</i>	0.055	0.010	0.071	0.011	0.049	0.009
<i>Female</i>	0.478	0.021	0.485	0.022	0.474	0.022
<i>Ideology (5-point scale)</i>	2.537	0.043	2.652	0.046	2.642	0.044
<i>News interest (4-point scale)</i>	2.132	0.039	2.058	0.038	2.164	0.040
<i>Registered to vote</i>	0.861	0.015	0.874	0.014	0.874	0.014
<i>White</i>	0.740	0.019	0.803	0.017	0.786	0.018
<i>Black</i>	0.090	0.012	0.056	0.010	0.071	0.011
<i>Other race</i>	0.170	0.016	0.141	0.015	0.143	0.015
<i>Religious importance</i>	2.711	0.081	2.835	0.082	2.846	0.084
<i>Northeast</i>	0.182	0.017	0.223	0.018	0.216	0.018
<i>Midwest</i>	0.233	0.018	0.214	0.018	0.216	0.018
<i>South</i>	0.344	0.020	0.344	0.020	0.348	0.021
<i>West</i>	0.241	0.018	0.219	0.018	0.220	0.018
<i>Elite</i>	0.079	0.012	0.087	0.012	0.099	0.013
<i>Observations</i>	<i>N</i> = 544		<i>N</i> = 538		<i>N</i> = 533	

References

- Angrist, J. D., G. W. Imbens, and D. B. Rubin. 1996. "Identification of Causal Effects using Instrumental Variables." *Journal of the American Statistical Association* 91(434):444–455.
- Aronow, P. M., J. Baron, and L. Pinson. 2018 "Replication Data for: A Note on Dropping Experimental Subjects who Fail a Manipulation Check." <https://doi.org/10.7910/DVN/GXXYMH>, Harvard Dataverse, V1.
- Berinsky, A. J., M. F. Margolis, and M. W. Sances. 2014. "Separating the Shirkers from the Workers? Making Sure Respondents Pay Attention on Self-Administered Surveys." *American Journal of Political Science* 58(3):739–753.
- Crawford, J. T., J. L. Brady, J. M. Pilanski, and H. Erny. 2013. "Differential Effects of Right-Wing Authoritarianism and Social Dominance Orientation on Political Candidate Support: The Moderating Role of Message Framing." *Journal of Social and Political Psychology* 1(1):5–28.
- De Oliveira, P., S. Guimond, and M. Dambun. 2012. "Power and Legitimizing Ideologies in Hierarchy-Enhancing versus Hierarchy-Attenuating Environments." *Political Psychology* 33(6):867–885.
- Gerber, A. S., and D. P. Green. 2012. *Field Experiments: Design, Analysis, and Interpretation*. WW Norton.
- Hauser, D. J., and N. Schwarz. 2016. "Attentive Turkers: MTurk Participants Perform Better on Online Attention Checks than do Subject Pool Participants." *Behavior Research Methods* 48(1):400–407.
- Hoffman, A. M., C. R. Agnew, L. E. VanderDrift, and R. Kulzick. 2013. "Norms, Diplomatic Alternatives, and the Social Psychology of War Support." *Journal of Conflict Resolution* 59(1):3–28.
- Lee, D. S. 2009. "Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects." *The Review of Economic Studies* 76(3):1071–1102.
- Maoz, I. 2006, 10. "The Effect of News Coverage Concerning the Opponents' Reaction to a Concession on its Evaluation in the Israeli–Palestinian Conflict." *The Harvard International Journal of Press/Politics* 11(4):70–88.

- Oppenheimer, D. M., T. Meyvis, and N. Davidenko. 2009. "Instructional Manipulation Checks: Detecting Satisficing to Increase Statistical Power." *Journal of Experimental Social Psychology* 45(4):867–872.
- Peer, E., J. Vosgerau, and A. Acquisti. 2014. "Reputation as a Sufficient Condition for Data Quality on Amazon Mechanical Turk." *Behavior Research Methods* 46(4):1023–1031.
- Press, D. G., S. D. Sagan, and B. A. Valentino. 2013. "Atomic Aversion: Experimental Evidence on Taboos, Traditions, and the Non-Use of Nuclear Weapons." *American Political Science Review* 107(01):188–206.
- Rubin, D. B. 1980. "Comment." *Journal of the American Statistical Association* 75(371):591–593.
- Small, D. A., J. S. Lerner, and B. Fischhoff. 2006. "Emotion Priming and Attributions for Terrorism: Americans' Reactions in a National Field Experiment." *Political Psychology* 27(2):289–298.
- Turner, J. 2007. "The Messenger Overwhelming the Message: Ideological Cues and Perceptions of Bias in Television News." *Political Behavior* 29(4):441–464.
- Wilson, T. D., E. Aronson, and K. Carlsmith. 2010. "The Art of Laboratory Experimentation." In *Handbook of Social Psychology*, 5th edn, edited by S. T. Fiske, D. T. Gilbert, G. Lindzey, and A. E. Jongsma, 51–81. Hoboken, NJ: Wiley.
- Zhang, J. L., and D. B. Rubin. 2003. "Estimation of Causal Effects via Principal Stratification When Some Outcomes are Truncated by 'Death.'" *Journal of Educational and Behavioral Statistics* 28(4):353–368.