



Voting behavior is unaffected by subtle linguistic cues: evidence from a psychologically authentic replication

ALAN S. GERBER

Charles C. & Dorathea S. Dilley Professor of Political Science, Yale University, New Haven, CT, USA

GREGORY A. HUBER *

Forst Family Professor of Political Science, Yale University, New Haven, CT, USA

ALBERT H. FANG 

Independent Scholar, New York, NY, USA

Abstract: Do small wording differences in message-based behavioral interventions have outsized effects on behavior? An influential initial study, examining this question in the domain of political behavior using two small-scale field experiments, argues that subtle linguistic cues in voter mobilization messages describing someone as a voter (noun) instead of one who votes (verb) dramatically increases turnout rates by activating a person's social identity as a voter. Two subsequent large-scale replication field experiments challenged this claim, finding no effect even in electorally competitive settings. However, these experiments may not have reproduced the psychological context needed to motivate behavioral change because they did not occur in highly competitive and highly salient electoral contexts. Addressing this major criticism, we conduct a large-scale, preregistered replication field experiment in the 2016 presidential election. We find no evidence that noun wording increases turnout compared to verb wording in this highly salient electoral context, even in competitive states.

Submitted 20 August 2020; revised 30 September 2020; accepted 9 October 2020

Introduction

An ongoing debate in behavioral science is about whether small differences in the wording of messages can have outsized effects on behavior. A leading example of this debate is scholarship about whether subtle linguistic cues

* Correspondence to: Forst Family Professor of Political Science, Yale University, New Haven, CT, USA. Email: gregory.huber@yale.edu

that link a desired behavior to a person's social identity are able to induce dramatic behavioral change by priming that identity. One operationalization of this concept is the argument that describing a person using a predicate noun (e.g., "to be a voter") emphasizes a behavior as an attribute of that person's social identity that can be claimed by engaging in that behavior. The theory argues that the use of a predicate noun, in contrast to describing a person's potential behavior using a verb (e.g., "to vote"), introduces a subtle linguistic cue that more clearly primes the behavior as related to one's identity and thus increases the likelihood that the person engages in it.

This argument, applied to the domain of political behavior, was initially advanced in an influential article in the *Proceedings of the National Academy of Sciences* by Bryan *et al.* (2011), who reported that priming a person's social identity as a voter using a predicate noun (instead of using a verb) in a 10-item Internet survey questionnaire completed either on the day before or the day of the election dramatically increased turnout by 11–14 percentage points in the 2008 general election in California and in the 2009 New Jersey gubernatorial election. If a subtle intervention of this sort can reliably produce such large behavioral effects, it would open the door to numerous promising opportunities in multiple domains to use policy to shape behavior. However, subsequent attempts to reproduce the result by Gerber *et al.* (2016, 2018) have found no difference between the effectiveness of noun and verb wording in increasing turnout, but there are multiple study parameters that vary across these studies.

In particular, in a response to the Gerber *et al.* (2016) study, three of the four original authors of the Bryan *et al.* (2011) study argued that replications must occur in contexts where the psychological phenomenon of interest could plausibly emerge such that a person's identity as a "voter" is salient enough to motivate behavioral change (Bryan *et al.*, 2016). Specifically, Bryan *et al.* (2016) argue that the psychological context would be reproduced only in electoral contexts that are both highly competitive and highly salient, and that the electoral settings from Gerber *et al.* (2016) (2014 primary elections in Michigan, Missouri, and Tennessee) did not meet these criteria. The electoral settings from the Gerber *et al.* (2018) study – which include contested gubernatorial elections in Kentucky, Louisiana, and Mississippi, as well as a contested mayoral election in Houston, all of which occurred in 2015 – are arguably open to the same critique.

This preregistered experiment, in which both the design of the experiment and the analysis of the data it produces are prespecified, addresses this major criticism and tests whether priming a registered voter's identity as a voter using a predicate noun (instead of a verb) leads to dramatic increases in actual turnout in a highly competitive and highly salient electoral context:

the 2016 presidential election. In addition, several other aspects of the experiment were designed in order to address potential questions about the robustness, replicability, and generalizability of the initial finding from Bryan *et al.* (2011) compared to subsequent work.

First, the treatments were delivered via an Internet survey on the day before Election Day in order to replicate the mode and timing of treatment delivery used by Bryan *et al.* (2011). Second, the experiment includes only subjects who were confirmed – by prematching against administrative voter records prior to the election – as being registered to vote in the 2016 presidential election. This eliminates individuals who could not vote in certain states. Third, subjects were also restricted to exclude confirmed registrants who voted by mail or absentee in 2014 in order to focus on registered voters for whom receiving any prime about their social identity as a voter from an Internet survey on the day of or the day before the election could plausibly change whether they vote. Fourth, the experiment included subjects from the two states examined by Bryan *et al.* (2011) – California and New Jersey – as well as from five other states: Connecticut, Michigan, New York, Ohio, and Pennsylvania. The seven states included in this experiment varied in terms of how competitive the presidential election was expected to be *ex ante*, which allows us to assess the robustness of noun wording effects across varying state-specific contexts of electoral competitiveness in an election that was both highly salient and highly competitive at the national level. Table 1 summarizes the key conditions that Bryan *et al.* (2016) argue are necessary to create the psychological context for noun wording to have a greater effect than verb wording on turnout and how they are satisfied by the present study and prior replication studies.

Study design

Subjects

The experiment was conducted during the 2016 presidential election in November. The 2219 subjects in our field experiment were US citizens from seven states (California, Connecticut, Michigan, New York, New Jersey, Ohio, and Pennsylvania) who were recruited from an online survey panel administered by YouGov and who were confirmed to be registered to vote in the November 2016 election prior to the experiment.¹

We focus on confirmed registrants from these seven states for the reasons described earlier. We restricted the total share of subjects from California

¹ A third-party vendor confirmed that subjects in the experiments were registered voters prior to the experiment by checking against state voter files.

Table 1. Conditions that Bryan *et al.* (2016) argue are necessary to create the psychological context for noun wording to be more effective than verb wording and how this and prior replication studies satisfy these conditions.

	Gerber <i>et al.</i> (2016)	Gerber <i>et al.</i> (2018)	This study
Study setting	2014 primary elections in MI, MO, and TN	2015 gubernatorial elections in KY, LA, and MS and mayoral election in Houston, TX	2016 presidential election; 7 states: CA, CT, MI, NY, NJ, OH, and PA
Highly salient elections? (“High-profile elections that received substantial public attention”)	No	High salience locally	High salience at national and state levels
Contested elections?	No	Yes	Yes
Competitive elections?	Not competitive	More competitive, both <i>ex ante</i> and <i>ex post</i>	Highly competitive nationwide, and includes both high- and low-competitive states

and Ohio (combined) to 30% due to concerns that the incidence of early voting in these states is high; the remaining 70% of subjects were recruited from the other five states (CT, MI, NJ, NY, and PA). The breakdown of the number of subjects recruited by state and treatment condition is reported in Table 2.

Importantly, we excluded registrants who voted by mail or absentee in 2014 because these subgroups were the least likely to be mobilized to vote in person on Election Day by an intervention administered on the day before Election Day. The subject pool included registrants who voted on Election Day, voted early, or did not vote at all in 2014.²

To replicate a key detail of the subject recruitment procedure from the experiments from study 3 by Bryan *et al.* (2011), subjects were recruited into our field experiment on the day before Election Day.

Treatments and randomization procedure

Subjects were randomly assigned to receive either a 10-item questionnaire using noun wording (“voter”), a 10-item questionnaire using verb wording that refers to the act of voting as a behavior (“voting/to vote”) or a placebo

² We included those who voted early in 2014 because we were unable to distinguish between those who voted early and those who voted on Election Day in 2014.

Table 2. Nonparametric estimates of turnout rates in the 2016 general election and differences in turnout rates between treatment arms.

Sample definition	Turnout rates									Difference in proportions					
	Placebo condition			Noun condition			Verb condition			Noun – placebo		Verb – placebo		Noun – verb	
	Prop	SE	<i>n</i>	Prop	SE	<i>n</i>	Prop	SE	<i>n</i>	Estimate	SE	Estimate	SE	Estimate	SE
Entire sample	0.778	0.018	541	0.759	0.015	850	0.785	0.014	828	-0.019	0.023	0.007	0.023	-0.026	0.020
California subsample	0.696	0.048	92	0.636	0.039	151	0.734	0.037	139	-0.060	0.062	0.038	0.061	-0.098	0.054
Connecticut subsample	0.759	0.079	29	0.806	0.066	36	0.690	0.071	42	0.047	0.103	-0.068	0.107	0.115	0.097
Michigan subsample	0.873	0.042	63	0.745	0.043	102	0.815	0.040	92	-0.128	0.060	-0.058	0.058	-0.070	0.059
New Jersey subsample	0.827	0.052	52	0.790	0.041	100	0.789	0.042	95	-0.037	0.066	-0.037	0.067	0.001	0.058
New York subsample	0.756	0.039	119	0.719	0.032	196	0.793	0.030	179	-0.037	0.051	0.037	0.050	-0.074	0.044
Ohio subsample	0.803	0.047	71	0.895	0.030	105	0.788	0.038	118	0.092	0.056	-0.015	0.060	0.107	0.048
Pennsylvania subsample	0.783	0.038	115	0.812	0.031	160	0.822	0.030	163	0.030	0.049	0.039	0.049	-0.010	0.043
High electoral competition states (MI, OH, and PA)	0.811	0.025	249	0.817	0.020	367	0.810	0.020	373	0.006	0.032	-0.002	0.032	0.008	0.029
Low electoral competition states (CA, CT, NJ, and NY)	0.750	0.025	292	0.714	0.021	483	0.765	0.020	455	-0.036	0.033	0.015	0.032	-0.051	0.029

condition asking how often the subject went to different retail establishments in the past week.

The full text of each treatment script is presented in the online Supplementary Information. The noun and verb treatment scripts are identical to those used in the study by Gerber *et al.* (2018), and are nearly identical to those used in study 3 by Bryan *et al.* (2011) and Gerber *et al.* (2016), where the minor differences do not alter the substantive meaning of the questions or the psychological interpretation of either treatment.

Treatment scripts were delivered using an Internet-based survey, replicating the treatment delivery mode used in study 3 by Bryan *et al.* (2011) and in Gerber *et al.* (2018).

The probability of assignment to each experimental condition was 37.5% for the noun condition, 37.5% for the verb condition, and 25.0% for the placebo condition. Respondents in the panel were enrolled in the experiment and assigned to a treatment arm immediately upon providing informed consent. We verify that the randomization is valid using randomization inference (see Supplementary Appendix) and present balance tables in Tables S1 and S2.

Outcomes

Our outcome measure is turnout in the 2016 general election, a behavior measured using voter files. Turnout data were supplied by the vendor, who located subjects in the voter files by matching on full name, address, year or date of birth, and gender. The turnout variable is coded 1 if the subject voted in the 2016 general election and 0 otherwise. This coding procedure is standard in the field experimental literature on voter mobilization. Subjects who cannot be located in state voter files after the election are coded as having not voted (i.e., outcome equals 0) to avoid introducing post-treatment bias that arises from differential attrition across treatment conditions.³

Results

Following our preanalysis plan, our primary analysis assesses the effect of the noun condition on turnout relative to assignment to the verb condition. In addition, we assess the effectiveness of the noun and verb conditions on turnout, both relative to the placebo condition.

³ Of the 2219 subjects, 18.6% could not be located in state voter files (these rates are 18.9% in the placebo condition, 19.3% in the noun condition, and 17.8% in the verb condition). Treatment assignment is not prognostic of whether the subject is located in voter files ($\chi^2(2) = 0.68$, $p = 0.71$).

Table 3. Regression estimates of the effect of noun and verb treatments on turnout in the 2016 general election.

	(1)	(2)
Voter (noun) treatment (1 = yes, 0 = no)	-0.019 (0.023)	-0.003 (0.019)
Voting (verb) treatment (1 = yes, 0 = no)	0.007 (0.023)	0.015 (0.019)
Constant	0.778*** (0.018)	0.312** (0.134)
Difference: noun – verb	-0.026 (0.020)	-0.018 (0.017)
P-values, one-tailed <i>t</i> -tests:		
H ₀ : noun – verb = 0; H _a : noun – verb > 0	0.900	0.859
H ₀ : noun – placebo = 0; H _a : noun – placebo > 0	0.800	0.559
H ₀ : verb – placebo = 0; H _a : verb – placebo > 0	0.384	0.211
With covariates?	No	Yes
Observations	2219	2219

p* < 0.1, *p* < 0.05, ****p* < 0.01, two-tailed *t*-tests unless otherwise specified.

The outcome variable is turnout in the 2016 general election (1 = yes, 0 = no). Covariates included in the covariate adjusted specification include age, age squared divided by 100, gender, race, education, survey date, party identification, ideology, past turnout in the 2016 primary and presidential primary elections and state fixed effects.

We begin in [Table 2](#) by presenting mean turnout rates by treatment arm and nonparametric estimates of differences in turnout rates between arms (noun versus verb, noun versus placebo, and verb versus placebo). In [Table 3](#), we supplement this nonparametric estimation with regression analysis, which allows us to test the robustness of our findings to the inclusion of individual-level covariates and state-level fixed effects. In particular, we estimate the following equation using ordinary least squares:

$$Y_i = \alpha + \beta_1 N_i + \beta_2 V_i + \gamma X_i + \varepsilon_i$$

where Y_i is turnout in the 2016 general election (1 = yes, 0 = otherwise); N_i is assignment to the noun condition (1 = yes, 0 = no); V_i is assignment to the verb condition (1 = yes, 0 = no); X_i is a vector of pretreatment covariates that include subjects' demographic characteristics (age, age squared divided by 100, gender, race, education, party identification, ideology, survey date, and past turnout in the 2016 primary and presidential primary elections) and state fixed effects; and ε_i is the error term. Pretreatment covariates are provided by YouGov. We estimate standard errors using the conservative Neyman estimator. The bottoms of the

regression tables report estimates of differences in effects across conditions and formal statistical tests of these differences. In order to test whether the noun condition is more effective at increasing turnout than the verb condition, we test the null hypothesis that $\beta_1 - \beta_2 = 0$ and calculate p-values and 95% confidence intervals (CIs) for a one-sided test (for an alternative hypothesis that $\beta_1 - \beta_2 > 0$). We additionally assess whether the noun condition is more effective as compared to the placebo and test the null hypothesis that $\beta = 0$ and calculate p-values and 95% CIs for a one-sided test (for an alternative hypothesis that $\beta_1 > 0$).

In all cases, results are similar for the nonparametric and regression analyses, and so we focus primarily on the nonparametric results shown in Table 2 for ease of presentation.

Pooled analysis

Nonparametric estimates pooling across all states appear in the first row of Table 2. A total of 77.8% of registrants in the placebo condition voted compared to 75.9% in the noun condition and 78.5% in the verb condition. Focusing on the difference between the noun and verb conditions, the last column of Table 2 shows that those in the noun condition were 2.6 percentage points *less* likely to vote than in the verb condition, although this difference is not statistically distinguishable from 0 ($z = -1.280$, $p = 0.899$, $b = -0.062$, 95% CI = $-0.062, 1$). The 95% CI therefore excludes the positive estimates of 11–14 percentage points reported in the field experimental studies in Bryan *et al.* (2011). Additionally, those in the noun condition were 1.9 percentage points less likely to vote than those who received no mobilization effort in the placebo condition, but this difference is not statistically significant ($z = -0.838$, $p = 0.797$, $b = -0.046$, 95% CI = $-0.057, 1$).

The regression estimates reported in Table 3 are highly similar. The column (1) specification, without covariates, matches the nonparametric results ($b(2216) = -0.026$, $p = 0.900$, $d = -0.031$, 95% CI = $-0.060, 1$). In the column (2) specification that incorporates state fixed effects and individual-level covariates, we continue to estimate that the noun treatment is less effective than the verb treatment at inducing turnout ($b(2180) = -0.018$, $p = 0.859$, $d = -0.026$, 95% CI = $-0.046, 1$). Notably, the noun treatment is also no more effective than the placebo message ($b(2180) = -0.003$, $p = 0.559$, $d = -0.004$, 95% CI = $-0.034, 1$).

Overall, these results provide little evidence that the noun treatment is more effective than the verb treatment at causing people to vote. We consistently estimate that the noun treatment is less effective than the verb treatment at

causing voting; although these differences are not statistically significant, they are relatively precisely estimated and exclude large positive effects. Nor is the noun message more effective than the placebo message.

Robustness: by state and by electoral competitiveness

In accordance with our preanalysis plan, we also examined differences by state and by *ex ante* electoral competitiveness.⁴ Subsetting by state necessarily reduces the sample sizes for different comparisons, and as such will increase sampling variability and the imprecision of our estimates. Per Table 2, in four states we estimate that the noun treatment is less effective than the verb treatment (California (estimate = -0.098, $z = -1.809$, $p = 0.963$, $h = -0.212$, 95% CI = -0.187, 1), Michigan (estimate = -0.070, $z = -1.185$, $p = 0.880$, $h = -0.170$, 95% CI = -0.167, 1), New York (estimate = -0.074, $z = -1.675$, $p = 0.952$, $h = -0.173$, 95% CI = -0.146, 1) and Pennsylvania (estimate = -0.010, $z = -0.223$, $p = 0.588$, $h = -0.025$, 95% CI = -0.080, 1)) and in three states it is more effective (Connecticut (estimate = 0.115, $z = 1.184$, $p = 0.123$, $h = 0.267$, 95% CI = -0.045, 1), New Jersey (estimate = 0.001, $z = 0.009$, $p = 0.496$, $h = 0.001$, 95% CI = -0.096, 1) and Ohio (estimate = 0.107, $z = 2.229$, $p = 0.015$, $h = 0.297$, 95% CI = 0.028, 1)). Only one of these nonparametric estimates is statistically distinguishable from 0 at test size $\alpha = 0.05$ given the smaller sample sizes.

Consistent with the greater role of sampling variability in smaller samples, the covariate adjusted regressions in Table S3 show that the largest estimated differences between the noun and verb conditions – both positive and negative differences – are attenuated in models with covariates and none are statistically significant. For example, in Connecticut, the 11.5 percentage point greater turnout among those in the noun rather than verb condition ($b(104) = 0.115$, $p = 0.125$, $d = 0.131$, 95% CI = -0.050, 1) is reduced to 3.6 percentage points when adjusting for pretreatment covariates ($b(79) = 0.036$, $p = 0.353$, $d = 0.043$, 95% CI = -0.123, 1), and the unadjusted estimate of 10.7 percentage point greater turnout in Ohio ($b(291) = 0.107$, $p = 0.017$, $d = 0.143$, 95%

⁴ We also assessed whether the main results were sensitive to restricting the sample to those who viewed and completed all of the survey items from their assigned treatment arm. While doing so may have introduced post-treatment bias if treatment receipt was affected by treatment assignment, we nonetheless present this preregistered analysis here as a footnote to allay potential concerns about the sensitivity of our main results to compliance with assigned treatment. This analysis, which appears in Tables S5 and S6, produces estimates that the noun treatment is less effective than the verb treatment by between 1.8 percentage points ($b(2175) = -0.018$, $p = 0.862$, $d = -0.032$, 95% CI = -0.046, 1) and 2.7 percentage points ($b(2211) = -0.027$, $p = 0.904$, $d = -0.026$, 95% CI = -0.060, 1), although again none of these estimates are statistically distinguishable from 0.

CI = 0.024, 1) is reduced to 4.9 percentage points with covariate adjustment ($b(263) = 0.049$, $p = 0.109$, $d = 0.083$, 95% CI = -0.017, 1). Similarly, the turnout rate in the noun condition is 9.8 percentage points lower than in the verb condition in California ($b(379) = -0.098$, $p = 0.963$, $d = -0.106$, 95% CI = -0.188, 1), but this effect is reduced to 3.3 percentage points with covariate adjustment ($b(350) = -0.033$, $p = 0.787$, $d = -0.047$, 95% CI = -0.101, 1). There is no state, therefore, in which there is statistically significant evidence that the noun condition is more effective than the verb treatment.

In light of the possibility that the noun treatment would be effective only in those competitive electoral environments where one's identity as a voter were potentially meaningful, we also present results by whether a state was deemed competitive in the 2016 election. In three states – Michigan, Ohio and Pennsylvania – pre-election forecasts led us to believe that the race would be close, and, in fact, those races were close, producing unexpected Republican victories in Michigan and Pennsylvania. Per Table 2, in these three states, turnout is 0.8 percentage points higher in the noun than verb conditions, but this estimate is not statistically significant ($z = 0.272$, $p = 0.393$, $h = 0.020$, 95% CI = -0.039, 1). In Table S4, where we adjust for state fixed effects and pretreatment covariates, this difference remains close to 0, and switches sign to -0.5 percentage points ($b(954) = -0.005$, $p = 0.584$, $d = -0.008$, 95% CI = -0.042, 1), which is again not significant. We note that, in all specifications, in the less competitive states the estimated effect of the noun treatment is to reduce turnout compared to the verb treatment. These estimates range from -3.0 percentage points ($b(1194) = -0.030$, $p = 0.886$, $d = -0.039$, 95% CI = -0.070, 1) to -5.1 percentage points ($b(1227) = -0.051$, $p = 0.961$, $d = -0.058$, 95% CI = -0.098, 1), but again none are significant.

Discussion

The promise for behavioral public policy of a psychological theory that claims subtle linguistic interventions can make salient a feature of one's identity and therefore change behavior is transparent. Rather than seeking to persuade or cajole, policymakers can seek to harness a person's own sense of self to encourage desirable behavior. In the domain of politics, this idea is particularly exciting, and an initial and influential study by Bryan *et al.* (2011) provided promising evidence that the use of noun language could prime individual's identities "as voters" to increase political participation in comparison to similar language using verbs ("voting"). Subsequent replication attempts did not yield similarly promising evidence (Gerber *et al.*, 2016, 2018), but they differed in ways that a subset of the authors of the original study argued were consequential (Bryan *et al.*, 2016).

Table 4. Summary of experimental studies testing noun wording (versus verb wording) on turnout.

Study	Treatment arms	Treatment delivery	Treatment timing	Electoral context (see Table 1 for more details)	Geographic restrictions	Screening criteria	Total sample size	Sample size in noun and verb groups	Outcome variables	Verb turnout rate	Reported effect on turnout (noun – verb)
Bryan <i>et al.</i> (2011), study 2	Noun block, verb block	Internet survey	On or before Election Day	November 2008 election	California	Self-reported as registered to vote, self-reported had not already voted (e.g., by mail) in election at time of survey, native English speakers, recruited via university-administered online participant pool on social networking site	88	88	Turnout as measured using matched voter file records	0.82	0.14
Bryan <i>et al.</i> (2011), study 3	Noun block, verb block	Internet survey	On or before Election Day	November 2009 election	New Jersey	Members of a randomly sampled and nationally representative panel administered by Knowledge Networks who were New Jersey residents, self-reported as registered to vote in New Jersey, self-reported not yet having voted at the time of the study	214	214	Turnout as measured using matched voter file records	0.79	0.11

Table 4. (Cont.)

Study	Treatment arms	Treatment delivery	Treatment timing	Electoral context (see Table 1 for more details)	Geographic restrictions	Screening criteria	Total sample size	Sample size in noun and verb groups	Outcome variables	Verb turnout rate	Reported effect on turnout (noun – verb)
Gerber <i>et al.</i> (2016)	Noun block, verb block, standard GOTV message	Telephone survey	4 days before and including Election Day	2014 primary elections	Michigan, Missouri, Tennessee	Registered voters with valid voter file records who could be contacted by phone	11,099	4468	Turnout as measured using matched voter file records	0.31	–0.01 (NS)
Gerber <i>et al.</i> (2018)	Noun block, verb block	Internet survey	4 days before and including Election Day	November 2015 election (gubernatorial and mayoral)	Gubernatorial elections (Kentucky, Louisiana, Mississippi); mayoral election (Houston, TX)	Recruited from online survey panels (SSI and YouGov). SSI: Self-reported residents of electoral jurisdiction, self-reported to be 18+, registered to vote, name provided (to be matched to voter file) YouGov: Self-reported resident of target jurisdictions, prematched to voter file to restrict to registered voters	3078	3078	Turnout as measured using matched voter file records	0.44	–0.02 (NS)

This study	Noun block, verb block, placebo	Internet survey	On or before Election Day	November 2016 presidential election	California, Connecticut, Michigan, New York, New Jersey, Ohio, Pennsylvania	Recruited from YouGov's online survey panel, confirmed to be registered to vote in November 2016 election before experiment, excluding registrants who voted by mail or absentee in 2014 (as they are least likely to be mobilized to vote in person on Election Day)	2219	1678	Turnout as measured using matched voter file records	0.79	-0.03 (NS)
------------	---------------------------------	-----------------	---------------------------	-------------------------------------	---	---	------	------	--	------	------------

GOTV = get out the vote; NS = nonsignificant; SSI = Survey Sampling International.

In light of this ongoing controversy, we provide novel evidence from a large-scale preregistered field experiment conducted during the 2016 presidential election that directly addresses these differences between the initial study and prior replication studies. Our experiment yields little evidence that priming subjects' identities as a voter by using noun language increases turnout compared to verb language, even in the salient 2016 presidential election and in competitive states. Moreover, neither treatment appears effective relative to a placebo message without any political content. In prespecified regression specifications controlling for pretreatment covariates, the point estimates from these studies for a noun versus verb treatment comparison are generally negative, and 95% CIs exclude the effects reported in Bryan *et al.* (2011). These results therefore imply both that those original results may have reflected sampling variability and that the estimates from different contexts (and differences in treatment delivery) reported in Gerber *et al.* (2016, 2018) are more representative of treatment effects even in competitive presidential contexts and with a design more closely mirroring the keystone study.

Overall, these results are disappointing, because they reveal that despite the potential promise of using subtle linguistic manipulations to encourage the prosocial behavior of voting, these messages appear largely ineffective. These messages, despite being longer and more complex than other mobilization messages (and therefore more costly to deliver), are not effective in promoting voting vis-à-vis one another or a nonpolitical placebo message. Additionally, the evidence undercuts the value of the theoretical perspective for efforts to encourage voting. More generally, the results reveal the importance of sustained and careful replication of prior work. In light of criticism of the study reported in Gerber *et al.* (2016) on the grounds that it was not an appropriate replication, this study addresses the arguments raised in Bryan *et al.* (2016) by more carefully recreating the context and method of treatment delivery in the original study. In this regard, and as Table 4 summarizes, careful comparison across studies and sustained efforts that help to rule out potential design-related sources of differences in results are shown, in this case, to lead to better and more dispositive evidence.

Thus, across this study and two earlier similar studies (that differ in the important features identified by Bryan *et al.* 2016), the weight of the evidence is that noun treatments are not more effective than verb language at increasing political participation.

Supplementary Material

To view supplementary material for this article, please visit <https://doi.org/10.1017/bpp.2020.57>.

References

- Bryan, C.J., G.M. Walton and C.S. Dweck (2016), 'Psychologically authentic versus inauthentic replication attempts', *Proceedings of the National Academy of Sciences of the United States of America*, **113**(43): E6548.
- Bryan, C.J., G.M. Walton, T. Rogers and C.S. Dweck (2011), 'Motivating voter turnout by invoking the self', *Proceedings of the National Academy of Sciences of the United States of America*, **108**(31): 12653–12656.
- Gerber, A.S., G.A. Huber, D.R. Biggers and D.J. Hendry (2016), 'A field experiment shows that subtle linguistic cues might not affect voter behavior', *Proceedings of the National Academy of Sciences of the United States of America*, **113**(26): 7112–7117.
- Gerber, A.S., G.A. Huber and A.H. 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.