Replication Proposal - Berlinski et al. (2023)

Andrew J. Liang

2026-02-17

We propose to replicate a prominent study showing that exposure to misinformation from the political elite decreases voters' trust in elections, and that fact-checking might be ineffective in combating the impact of misinformation regarding election fraud for voters (Berlinski et al. 2023). In that study, 4,283 respondents from YouGov were randomly assigned to either view four food-related tweets (the control condition), a random subset of four out of eight tweets from prominent Republican figures claiming election fraud/meddling in the 2018 midterm elections (the low treatment condition), all eight of the election fraud tweets (high treatment condition), or four random election fraud tweets plus four fact-checking tweets from prominent news sources such as the *New York Times* (the fact-checking treatment condition) in a between-subjects design. The authors found that exposure to social media posts from prominent political figures suggesting that election fraud may have occurred in the 2018 midterm elections decreased voters' confidence in election integrity, and that even if these posts were supplemented by a fact-checking source, that decrease in trust persisted. The authors also found that the fact-checking tweets were not substantially effective at reducing the effects of misinformation on voter confidence, compared to individuals who saw the four misinformation tweets without fact-checking. We aim to introduce a similar experiment across a 8,000-respondent sample from Rep Data between November and December 2026, after the 2026 midterm elections have concluded.

Replication of this study is important for three reasons:

1. Although concerns over whether voters are confident that elections are being administered fairly have long existed in the literature (Craig et al. 2006; Minnite 2010), trust in elections has dropped precipitously within the last decade (Jaffe et al. 2024; Stewart 2022). Claims of election fraud have increased significantly in the years after the 2018 midterm election, particularly after the presidential election in 2020 when Donald Trump claimed that the election was stolen from him; these claims may carry long-term consequences for elections officials and the public (Barrett, Corasaniti, and Fausset 2026; Winter and Goudsward 2026). When political elites are the ones voicing these claims, voters may adjust their perceptions of elections accordingly (Clayton et al. 2021; Agadjanian 2021). As a result, media organizations and elections officials have begun fact-checking campaigns against misinformation in the hopes of regaining confidence in elections among voters (Coppock et al. 2023).

   The question that remains, however, is to what extent does fact-checking have an impact on combating misinformation. While Berlinski et al. (2023) found that fact-checking social media posts from media organizations did not appear to substantially impact the effects of misinformation on voter confidence, recent research has indicated that "prebunking" messages may be able to positively shape voters' trust in elections (Carey et al. 2025), and that elections officials may also be more successful in instilling confidence in elections (Gaudette et al. 2025; Lockhart et al. 2024). This effectiveness also appears to translate to official messaging transmitted through social media (Suttmann-Lea and Merivaki 2023), so what could cause this difference in outcomes? One plausible reason is that media organizations are now producing shorter, "bite-sized" news articles with very little substantive information (Trexler 2026). This, coupled with Americans' growing distrust of the media (Ladd 2012), suggests that voters use the source of the fact-checker as a signal of the message's persuasiveness. Replicating Berlinski et al. (2023) would help us confirm whether the source of the fact-checking matters in increasing voter confidence; since the original study focused exclusively on social media posts, similar results could indicate the possibility that a fact-checker's authority matters in their persuasiveness, and that social media posts from media organizations may be ineffective as fact-checking supplements to misinformation.

2. The authors conducted their own power analysis and found that while the estimates for the low and high treatment doses were sufficiently powered, the estimate for the fact-checking dose was not. This is especially concerning since a key focus of the paper was whether fact-checking could counteract the effects of misinformation on voters' confidence in elections. After replicating the authors' power analysis, we show that the necessary sample size for all three estimates to be sufficiently powered is approximately 8,000 respondents. We believe that Rep Data's ability to provide both high-quantity *and* high-quality respondents can help us overcome issues of statistical power for our main estimands.

3. 2026 is a midterm election year, similar to the 2018 midterm election studied in the original paper. However, the setting is different in that we now study how voters respond to misinformation during elections after notable elections-related events like the January 6th Capitol attack and the Dominion-Fox News lawsuit have concluded. This gives us a chance to see whether these events, which negatively impacted some voters' perceptions of election integrity across party lines (Herron 2023), can explain why we observe heterogeneous treatment effects across party lines.

Replication

We seek to replicate Berlinski et al. (2023) with a nationally-representative, 8,000-individual sample from Rep Data. We would use an identical design to that of the original study, combining the pre-treatment questions asked in Wave 1 into a single wave of survey questions with our treatment questions (as well as the questions in our novel survey experiments). Prior to introducing the treatment conditions, we would ask respondents about several characteristics typical of other public opinion surveys, such as their highest level of education, age, and ideological leaning.

We will randomly assign respondents to one of four conditions: a control condition where respondents are asked to view four food-related posts; a low-treatment condition where respondents are asked to view a random subset of four out of eight possible posts where a prominent political figure made a claim of election fraud; a high-treatment condition where respondents are asked to view all eight of the possible posts with election fraud, and a fact-checking condition similar to the low-treatment condition, but with an additional four posts describing a fact-checking article issued by a media organization, such as the *New York Times*. Our complete randomization should yield similar assignment groups as those shown in Table 1.

Table 1: Expected treatment assignment for Rep Data sample.

	Control	High Dose	Low Dose	Low Dose + Fact-Check Posts	Total
Number of Units	2000	2000	2000	2000	8000

We choose to reanalyze three estimands of the paper - the treatment effect of various exposure levels of election fraud on a composite measure of respondents' confidence towards elections, the difference in the treatment effects between the high treatment dose and the low treatment dose on the composite outcome, and the difference in the treatment effects between the low treatment dose and the fact-checking treatment dose. These estimands correspond to hypotheses H1a, H2a, RQ1a, H3a, and H4a in the original study, and to Table 2, Column 8 in the paper. We choose these estimands specifically because they can help us address some of the questions we posed in the introduction; confirming whether fact-checking social media posts are in fact ineffective at reducing the impact of misinformation would help us transition to the objective of our novel survey experiment - verifying whether the source of the fact-checking matters.

We present two options for the replication study based on a potential complication with the treatment conditions offered to respondents. We originally planned to update the social media posts used for each treatment condition so that they would accurately reflect the sentiment of prominent political figures during the 2026 election. The original study used posts from Twitter for the treatment, and two of then-President Donald Trump's tweets were used as the treatment dose in the original study. After the January 6th Capitol attacks, however, Trump was banned from Twitter and subsequently formed his own social media network, Truth Social, in 2022. Furthermore, since Elon Musk's acquisition of Twitter (and its renaming to X), some voters have chosen to migrate to other platforms like Bluesky Social or Truth Social. As a result, voters are now likely to have three different platforms from which they could choose to receive elections-related information. It is also likely, however, that voters are able to infer the purpose of our survey experiment if we were to provide posts from platforms other than X due to demand effects (Mummolo and Peterson 2019). As a result, we present two options for creating treatment conditions in the replication study:

1. We can proceed by using the exact same tweets in the original study for our treatment conditions.
2. We can select new posts for the treatment conditions solely from X for the reasons described above; we would gather these new tweets for our treatment conditions closer to the 2026 general election. We would probably gather eight posts along the same proportion of type of political figure as in the original study (two tweets from the President, one from the Republican Party, five from Senate officials/candidates).

Reanalysis

To construct the composite measure of confidence (the main outcome variable), the authors estimated a structural equation model to predict a standardized outcome variable for confidence towards elections based on seven preregistered factors. We used the `psych` and `lavaan` packages in R for this replication. While the coefficients of the structural equation model were slightly different from the results in the Online Appendix, we found that the constructed outcome variable had similar statistics to those reported in Table 1 of the original article.

We successfully replicated the results of Table 2, Column 8, in our analysis; Table 2 displays the estimates of the ordinary least squares models specified in the original article.[1] The results suggest that exposure to unverified voter fraud claims decreases confidence in elections, regardless of the level of treatment given. For example, voters who were shown four tweets about election fraud saw a 0.147 sample standard-deviation decrease in confidence towards elections compared to the control group. In addition, those who were shown fact-checking tweets after the four election fraud tweets saw an approximately 0.092 standard-deviation decrease in confidence towards elections compared to the control group, indicating that fact-checks did not entirely reinstill voters' confidence in elections. Since the authors only reported unadjusted estimates as part of their study, we also report covariate-adjusted estimates for our treatment effects in Table 2. We controlled for factors like gender, race, education, age, political knowledge, martial status, and political interest, interacting our treatment indicators in accordance with Lin (2013). Our results are similar to the unadjusted estimates of the treatment effects, so we plan to include these covariates in our replication (see the survey instrument for more details).

The authors also reported results of a power analysis in the Online Appendix, but only to test whether the treatment effects were large enough to achieve sufficient power. We successfully replicated the results of that analysis according to the procedure outlined as Online Appendix F; our results are illustrated in the left three graphs of Figure 1. The treatment effects for the low and high treatment conditions are sufficiently powered, but the effect for the fact-checking treatment condition is underpowered.

Since the effect of fact-checking was a key focus of the original paper (and still is a main focus for our novel survey experiment), we conducted our own power analysis using the `DeclareDesign` package (Blair et al. 2019) to understand the sample size we would hypothetically need for our observed estimate to have sufficient power. The right three graphs of Figure 1 show that while the study of the composite confidence score was sufficiently powered to detect the estimated treatment effects for the low and high treatment conditions, we would need approximately 7,300 respondents for the fact-checking treatment effect size to be sufficiently powered as well. **Because of this we seek a sample of 8,000 respondents for our replication.**

We would also like to test whether the effects are heterogeneous among voters who identify as Republicans, Democrats, and Independents. It has been well-established that Republicans and Democrats differ in both whether they believe election fraud has occurred (Fahey 2023), as well as their concern as to whether voters are able to cast their votes in elections (Huber et al. 2025). While the authors did test for heterogeneous treatment effects among party identification and approval of then-President Trump, their original hypotheses only suggested heterogeneous treatment effects between Republicans and non-Republicans; we believe that Republicans, Democrats, and Independents will react differently to both exposure to misinformation about election fraud as well as fact-checking against misinformation. We replicated the results of the authors' original analysis; our results are reported in the appendix as Figure 2. Notably, treatment effects were similar for Republicans and Independents across treatment conditions, whereas Democrats were less likely to see reduced trust in elections after exposure to unfounded claims. Because most of the effects for Democrats and Independents are not statistically significant, however, this analysis would also benefit from a larger sample size from Rep Data.

[1] While we only focused on the main outcome variable of the original study, we successfully replicated all of the tables in the main portion of the paper as well as the figures. The code used to replicate those tables can be found as `Replication.R` in the online GitHub repository for this proposal.

References

Agadjanian, Alexander. 2021. "When Do Partisans Stop Following the Leader?" *Political Communication* 38 (4): 351–69. https://doi.org/10.1080/10584609.2020.1772418.

Barrett, Devlin, Nick Corasaniti, and Richard Fausset. 2026. "Georgia Ballot Inquiry Originated from Election Denier in Trump White House." https://www.nytimes.com/2026/02/10/us/politics/fulton-county-kurt-olsen-fbi-search-2020-ballots.html; The New York Times.

Berlinski, Nicolas, Margaret Doyle, Andrew M. Guess, Gabrielle Levy, Benjamin Lyons, Jacob M. Montgomery, Brendan Nyhan, and Jason Reifler. 2023. "The Effects of Unsubstantiated Claims of Voter Fraud on Confidence in Elections." *Journal of Experimental Political Science* 10 (1): 34–49. https://doi.org/10.1017/XPS.2021.18.

Blair, Graeme, Jasper Cooper, Alexander Coppock, and Macartan Humphreys. 2019. "Declaring and Diagnosing Research Designs." *American Political Science Review* 113: 838–59. https://doi.org/10.1017/S0003055419000194.

Carey, John M., Brian Fogarty, MarÃ■lia Gehrke, Brendan Nyhan, and Jason Reifler. 2025. "Prebunking and Credible Source Corrections Increase Election Credibility: Evidence from the US and Brazil." *Science Advances* 11 (35): eadv3758. https://doi.org/10.1126/sciadv.adv3758.

Clayton, Katherine, Nicholas T. Davis, Brendan Nyhan, Ethan J. Porter, Timothy J. Ryan, and Thomas J. Wood. 2021. "Elite Rhetoric Can Undermine Democratic Norms." *Proceedings of the National Academy of Sciences of the United States of America* 118 (23): e2024125118. https://doi.org/10.1073/pnas.2024125118.

Coppock, Alexander, Kimberly Gross, Ethan Porter, Emily Thorston, and Thomas J. Wood. 2023. "Conceptual Replication of Four Key Findings about Factual Corrections and Misinformation During the 2020 US Election: Evidence from Panel-Survey Experiments." *Political Analysis* 53: 1328–41. https://doi.org/10.1017/S0007123422000631.

Craig, Stephen C., Michael D. Martinez, Jason Gainous, and James G. Kane. 2006. "Winners, Losers, and Election Context: Voter Responses to the 2000 Presidential Election." *Political Research Quarterly* 59 (4): 579–92. https://doi.org/10.1177/106591290605900407.

Fahey, James J. 2023. "The Big Lie: Expressive Responding and Misperceptions in the United States." *Journal of Experimental Political Science* 10 (2): 267–78. https://doi.org/10.1017/XPS.2022.33.

Gaudette, Jennifer, Seth J. Hill, Thad Kousser, Mackenzie Lockhart, and Mindy Romero. 2025. "Can Official Messaging on Trust in Elections Break Through Partisan Polarization?" *British Journal of Political Science* 55: e16. https://doi.org/10.1017/S0007123424000681.

Herron, Michael C. 2023. "Allegations Made Against Dominion Voting Systems and the 2020 Presidential Election in Wisconsin." *Election Law Journal* 22 (3): 247–67. https://doi.org/10.1089/elj.2022.0051.

Huber, Gregory A., John J. Cho, Scott E. Bokemper, Alan S. Gerber, William J. Brady, Killian McLoughlin, and Molly J. Crockett. 2025. "Disagreements about Threats to Electoral Integrity: Beliefs about the Severity and Frequency of Fraudulent, Uncounted, and Forgone Votes in the 2020 and 2024 Elections." *Political Behavior* 0 (0): 1–26. https://doi.org/10.1007/s11109-025-10066-y.

Jaffe, Jacob, Joseph R. Loffredo, Samuel Baltz, Alejandro Flores, and Charles Stewart. 2024. "Trust in the Count: Improving Voter Confidence with Post-Election Audits." *Public Opinion Quarterly* 88: 585–607. https://doi.org/10.1093/poq/nfae029.

Ladd, Jonathan M. 2012. *Why Americans Hate the Media and How It Matters*. Princeton University Press.

Lin, Winston. 2013. "Agnostic Notes on Regression Adjustments to Experimental Data: Reexamining Freedman's Critique." *The Annals of Applied Statistics* 7 (1): 295–318. https://doi.org/10.1214/12-AOAS583.

Lockhart, Mackenzie, Jennifer Gaudette, Seth J. Hill, Thad Kousser, Mindy Romero, and Laura Uribe. 2024. "Voters Distrust Delayed Election Results, but a Prebunking Message Inoculates Against Distrust." *PNAS Nexus* 3 (10): pgae414. https://doi.org/10.1093/pnasnexus/pgae414.

Minnite, Lorraine C. 2010. *The Myth of Voter Fraud*. Cornell University Press.

Mummolo, Jonathan, and Erik Peterson. 2019. "Demand Effects in Survey Experiments: An Empirical Assessment." *American Political Science Review* 113: 517–29. https://doi.org/10.1017/S0003055418000837.

Stewart, Charles. 2022. "Trust in Elections." *Daedalus* 151 (4): 234–53. https://doi.org/10.1162/daed_a_01953.

Suttmann-Lea, Mara, and Thessalia Merivaki. 2023. "The Impact of Voter Education on Voter Confidence: Evidence from the 2020 u.s. Presidential Election." *Election Law Journal: Rules, Politics, and Policy* 22 (2): 145–65. https://doi.org/10.1089/elj.2022.0055.

Trexler, Andrew. 2026. "The Unequal Challenge of Learning from Under-Informative News." *British Journal of Political Science* 56: e9. https://doi.org/10.1017/S0007123425101300.

Winter, Jana, and Andrew Goudsward. 2026. "FBI Searches Atlanta Election Office, Chasing Trump 2020 Vote Fraud Claims." https://www.reuters.com/world/us/fbi-executing-search-warrant-election-office-georgia-related-2020-vote-fox-news-2026-01-28/; Reuters.

Appendix: Reanalysis Results

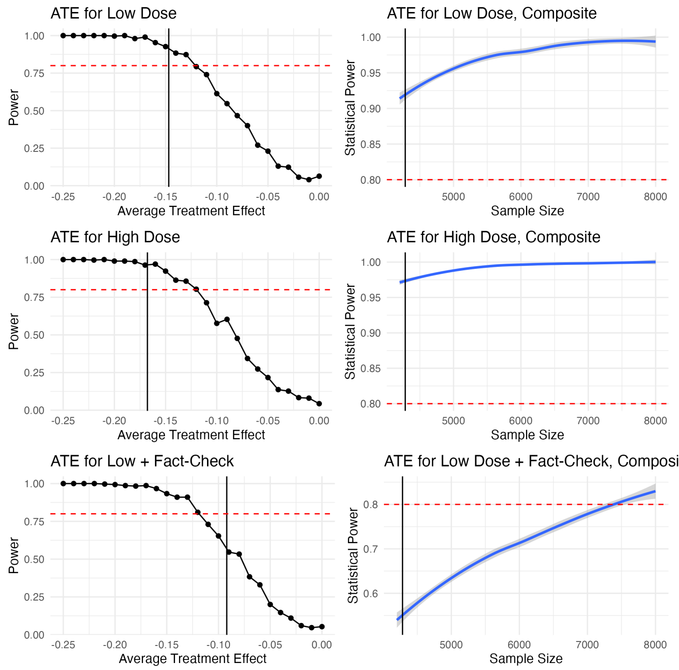

Figure 1: Replication of power analysis and sample sizes required for sufficiently powered effect sizes, by treatment condition. The dashed red line represents the conventional level of power (80%). The black line indicates the sample size of the original study (N = 4,283), or the estimated effect sizes for each treatment condition.

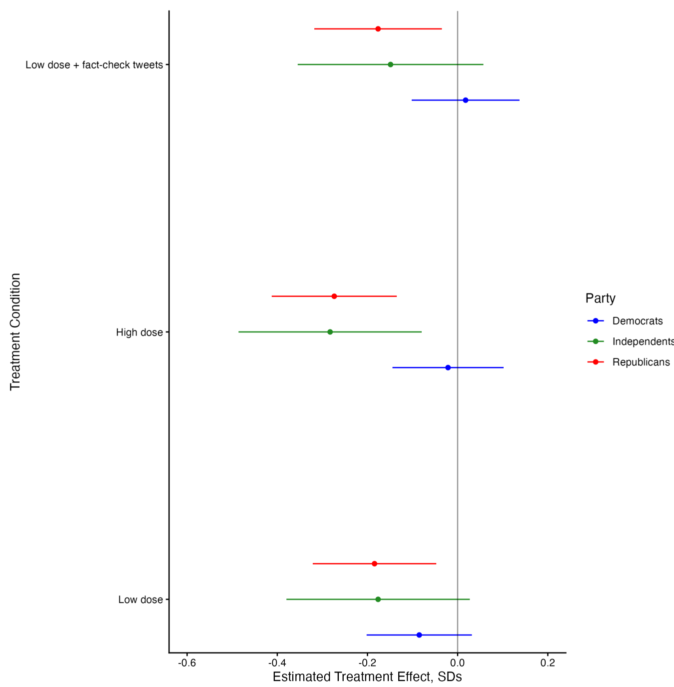

Figure 2: Replication of estimated heterogeneous treatment effects across party identification, originally reported as Figure 3(a) in the main paper. Estimates obtained using ordinary least squares regression; 95% confidence intervals constructed using HC2 robust standard errors.

Table 2: Replication Results of Table 2, Column 8 of Berlinski et al. (2023). Simple difference-in-means and covariate-adjusted estimates obtained using OLS regression. 95% Confidence Intervals calculated using HC2 robust standard errors.

	Simple Difference-in-Means			Covariate-Adjusted		
	Main Model	High - Low Dose	Low + Fact-Check - Low Dose	Main Model	High - Low Dose	Low + Fact-Check - Low Dose
Low dose (H1a)	-0.147^*					
	$[-0.230; -0.064]$					
High dose (H2a)	-0.168^*				-0.008	
	$[-0.253; -0.083]$				$[-0.091; 0.074]$	
Low dose + fact-check tweets (RQ1a)	-0.092^*					0.072
	$[-0.177; -0.007]$					$[-0.011; 0.155]$
High - Low Dose		0.076		-0.168^*		
		$[-0.011; 0.162]$		$[-0.251; -0.085]$		
Low + Fact-Check - Low Dose			0.055	-0.088^*		
			$[-0.029; 0.139]$	$[-0.172; -0.005]$		
Constant	0.102^*	-0.066^*	-0.045	0.105^*	-0.050	-0.052
	$[0.043; 0.161]$	$[-0.127; -0.004]$	$[-0.103; 0.013]$	$[0.048; 0.163]$	$[-0.107; 0.006]$	$[-0.110; 0.005]$
Num. obs.	4280	2141	2145	4273	2146	2141

* Null hypothesis value outside the confidence interval.