SHORT ARTICLE

Political Misinformation and Factual Corrections on the Facebook News Feed: Experimental Evidence

Ethan Porter, George Washington University
Thomas J. Wood, Ohio State University

As concerns about the spread of political misinformation have mounted, scholars have found that fact-checks can reduce the extent to which people believe misinformation. Whether this finding extends to social media is unclear. Social media is a high-choice environment in which the cognitive effort required to separate truth from fiction, individuals' penchant for select exposure, and motivated reasoning all may render fact-checks ineffective. Furthermore, large social media companies have largely refrained from permitting external researchers to administer experiments on their platforms. To investigate whether fact-checking can rebut misinformation on social media, we administer two experiments on large, nationally representative samples using a novel platform designed to mimic Facebook's news feed. We observe corrections having large effects on factual beliefs (.62 on a five-point scale, $p < .001$). While misinformation degrades accuracy, our results offer strong evidence that fact-checks increase accuracy, even when tested on realistic simulations of social media platforms.

The proliferation of misinformation represents a pressing policy issue, with potentially deleterious consequences for US democracy. In the months preceding the 2016 election, Facebook users encountered more fraudulent stories supporting the eventual winner than the eventual loser (Allcott and Gentzkow 2017). While the spread of misinformation on social media should not be exaggerated (Guess, Nyhan, and Reifler 2020), it also should not be dismissed. The possible political consequences of misinformation are wide-ranging in scope (Jamieson 2018), and access to empirically accurate information stands out as critical for democratic functioning (Einstein and Hochschild 2015).

Factual corrections offer one potential solution. Although earlier research indicated that corrections could not only be ineffective but "backfire" and increase inaccuracy (Nyhan and Reifler 2010), more recent findings have reached the opposite conclusion. Corrections can increase factual accuracy (Walter et al. 2020), even when the corrections target copartisans in periods of intense political competition (Wood and Porter 2019).

Yet whether fact-checks can increase belief accuracy in high-choice social media environments is unclear. Given the relationship between cognitive effort and ability to accurately identify misinformation (Pennycook and Rand 2019), social media users may avoid exerting the cognitive effort required to sift through fact-checks, instead choosing to engage with less demanding content. In line with theories of selective exposure (Stroud 2011), users could choose to avoid reading fact-checks that correct copartisans (Hameleers and van der Meer 2020). Alternatively, users could choose to consume fact-checks but reconcile them with their partisanship (Lodge and Taber 2013).

We present two preregistered experiments that advance our understanding of the effectiveness of factual corrections on social media. Both experiments were fielded on a platform engineered to resemble Facebook's news feed. While other

Ethan Porter (evporter@gwu.edu) is an assistant professor at George Washington University. Thomas J. Wood (wood.1080@osu.edu) is an assistant professor at Ohio State University.

The study was declared exempt by the Ohio State University Institutional Review Board (no. 2019E0967). Replication files are available in the *JOP* Dataverse (https://dataverse.harvard.edu/dataverse/jop). The empirical analysis has been successfully replicated by the *JOP* replication analyst. An appendix with supplementary material is available at https://doi.org/10.1086/719271.

Published online May 3, 2022.

 ooo

researchers have studied simulated news environments (Kalmoe et al. 2018), ours is among the first that we are aware of to study corrections and misinformation on a detailed simulation of the news feed (but see Bode and Vraga 2015). Prior researchers have studied learning on the news feed (Bode 2016), testing individual items such as headlines (Pennycook et al. 2020) and links (Vraga et al. 2016), while compelling subjects to consume misinformation, corrections, and similar material. Subjects enrolled in our experiments were exposed to news feeds that contained, at random, multiple items of misinformation, factual corrections, and nonpolitical placebo content. Distinct from participants in other studies on this topic, our subjects were free to read, and avoid reading, any material that they wished.

The stakes of these experiments are considerable, for social media platforms and US democracy. More Americans rely on Facebook for news than any other social media platform (Shearer and Matsa 2018). As prior scholars have emphasized (Delli Carpini and Keeter 1996), democracy depends on citizens receiving factually accurate information and being able to distinguish fact from fiction. If fact-checks can increase accuracy on social media, this would suggest that the challenges posed by misinformation to democracy can be curtailed. Alternatively, if fact-checks cannot increase accuracy on such a popular source of news and information, the relationship between empirical evidence and democracy would be at risk.

DESIGN

Upon entering our survey, participants were exposed to a social news feed modeled on the Facebook news feed.[1] On the first news feed, each participant was exposed to five stories; at random, 0–5 were "fake" stories, while the remaining were placebo stories. Next, participants were taken to a second news feed that, at random, featured 0–5 factual corrections. Subjects could only be randomly exposed to corrections on the second news feed if they had seen the related misinformation on the first feed. Randomization occurred at the claim level. For every tested fake story, subjects were in one of three conditions, in which they were exposed to (1) the misinformation and a survey item or (2) the misinformation, a factual correction, and a survey item or (3) the survey item only.

We did not compel subjects to read any items on either news feed. Subjects could have scrolled through both the first and second feeds without paying attention to the misinformation or corrective content and instead read the placebos or not read any material at all. Subjects were not incentivized for accuracy or instructed to read all material. To compare our fact-checks to those used on Facebook today, inspect figure 1.

The tested fake stories had actually circulated on Facebook, including false claims relating to immigrants and measles, climate activist Greta Thunberg, and President Trump. Corrections directly contradicted the false claims and offered subjects the opportunity to access a more detailed fact-check on an external website. Placebos included recipes for macaroni and cheese, a picture of a coffee shop, and a story about sign language in airports. The tested fact-checks resembled those typical of fact-checking organizations. The order of stories across news feeds was randomized. All participants then answered questions that measured their factual beliefs about all five items.

The second experiment investigated whether results from the first would be robust to changes in correction design and outcome measurement. In the first experiment, if participants were assigned to see a correction, on the second news feed the correction would appear above a facsimile of the original fake story. In the second experiment, on the second news feed the correction would appear above a version of the original fake story that had been blurred out. In the first experiment, participants evaluated each false story on a 1–5 agreement scale; in the second, they responded on a 1–5 truthfulness scale.[2] The survey items appear in the appendix. Our hypotheses and analysis strategy were preregistered.

RESULTS

Both experiments were administered by YouGov on US respondents. In the first experiment ($n = 5{,}000$), fielded October 24–November 1, 2019, subjects were surveyed and then matched to a sampling frame based on the 2016 American Community Survey. Experiment 2 ($n = 2{,}000$) was administered on December 2–4, 2019, and January 22–28, 2020, and also matched to the American Community Survey.

For both experiments, each five-point scale of agreement was modeled as a dependent quantity using linear regression, where the predictive variable was an experimental condition. Three conditions were used for each item: whether respondents saw the survey item only, whether they saw the uncorrected misinformation and survey item, or whether they saw the misinformation, a correction, and the survey item.

Corrections provoked large gains in accuracy.[3] Figure 2 displays results (regression tables are in the appendix). We

1. The platform was designed by Avaaz, a civil society group, for research and advocacy purposes and was modified to meet our specifications.

2. We relied on separate question wordings out of concern (Swire-Thompson, DeGutis, and Lazer 2020) that findings of backfire may be owed to measurement error stemming from question wording.

3. Table 3 displays effect size by Cohen's d, which our discussion of effect size relies on.

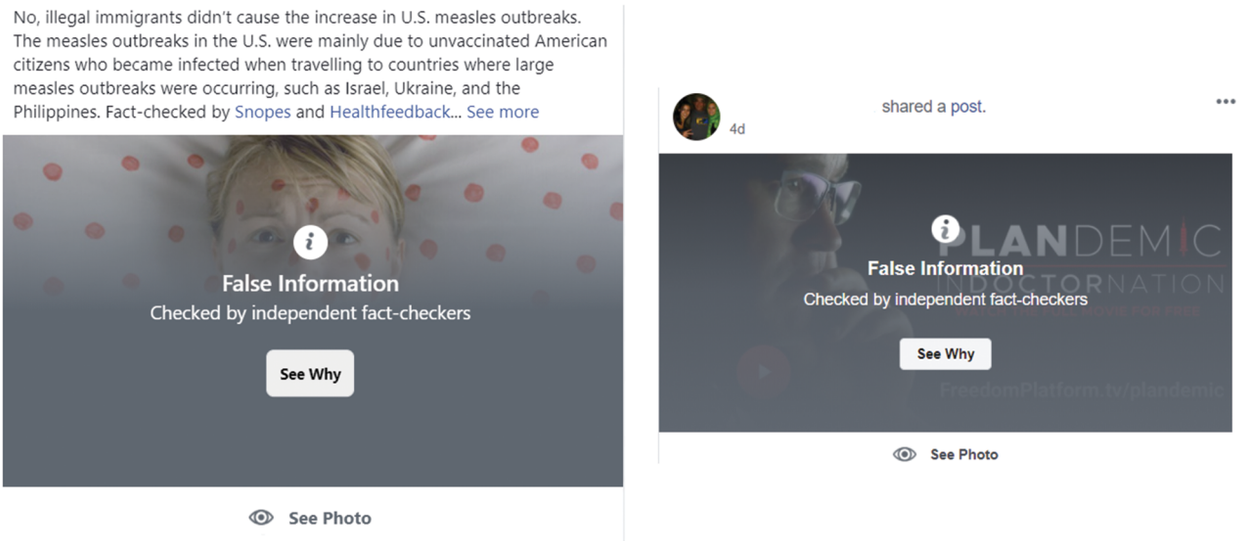

Figure 1. *Left*, fact-checked treatment from our studies; *right*, fact-checked item on Facebook in 2020

display item-by-item and overall effects. In experiment 1, the overall correction effect was .55 ($p < .001$) on a five-point scale in the direction of greater accuracy. In experiment 2, the overall correction effect was .79 ($p < .001$), again on a five-point scale. Averaging across experiments and weighting for sample size, the mean correction effect was .62 ($p < .001$).

Misinformation reduced factual accuracy. However, the effects of misinformation sans correction were smaller. In experiment 1, the average effect of being exposed to misinformation without a correction was −.12 ($p < .001$) on a five-point scale in the direction of reduced accuracy, while in experiment 2 it was −.14 ($p < .001$) on the same scale, again in the direction of reduced accuracy. Averaging across both experiments and weighting for sample size, the mean misinformation effect was −.13 ($p < .001$). In both experiments, the size of the correction effect was much larger than the size of the misinformation effect ($p < .001$). Fact-checks decrease false beliefs by larger amounts than misinformation, without corrections, increases them.

One way of encapsulating the magnitude of correction effects is to compare them to the differences among liberals and conservatives in the "items-only" condition, who saw neither the misinformation nor the correction. Averaging across experiments and issues, subjects on opposite ends of the ideological spectrum exhibited a .5 difference along our five-point scale of agreement. The effect of corrections on accuracy was thus greater than the mean difference between liberals and conservatives who had seen neither misinformation nor corrections. In line with recent work, we do not find conservatives to be more resistant to facts (Ryan and Aziz 2021). Figure 3 displays effects by party. Effects are mostly consistent, with some variation by issue. In the appendix, we present further heterogeneous effects by subgroup, finding no backfire.

DISCUSSION

Factual corrections can reduce belief in misinformation, even on realistic social media simulations. To be sure, fact-checks alone cannot eliminate the misinformation problem. As we show, when not followed by a factual correction, misinformation reduces factual accuracy. Scaling up factual corrections to match the scope of the misinformation challenge represents a very large task; other solutions will likely have to be used as well.

That being said, while prior research has shown that individuals respond to corrections by becoming more accurate, this study demonstrates the same is true in a high-choice social media environment. Such environments pose challenges to democracy (Prior 2007) and rebuttals to misinformation (Hameleers and van der Meer 2020). Our subjects could have scrolled through the news feeds without reading the fact-checks, read only the fact-checks that accorded with their partisanship, or relied on motivated reasoning to negate the fact-checks they read. Instead, on average, across partisan and ideological lines, subjects read fact-checks and became more accurate.

Separating fact from fiction remains difficult (Brashier and Marsh 2020). Yet even on a platform that approximates Facebook's news feed, in which subjects were presented with politically incongruent fact-checks, corrections increased accuracy. Although conservatives and older Americans may consume misinformation at comparatively higher rates (Guess, Nagler, and Tucker 2019), both groups responded to corrections

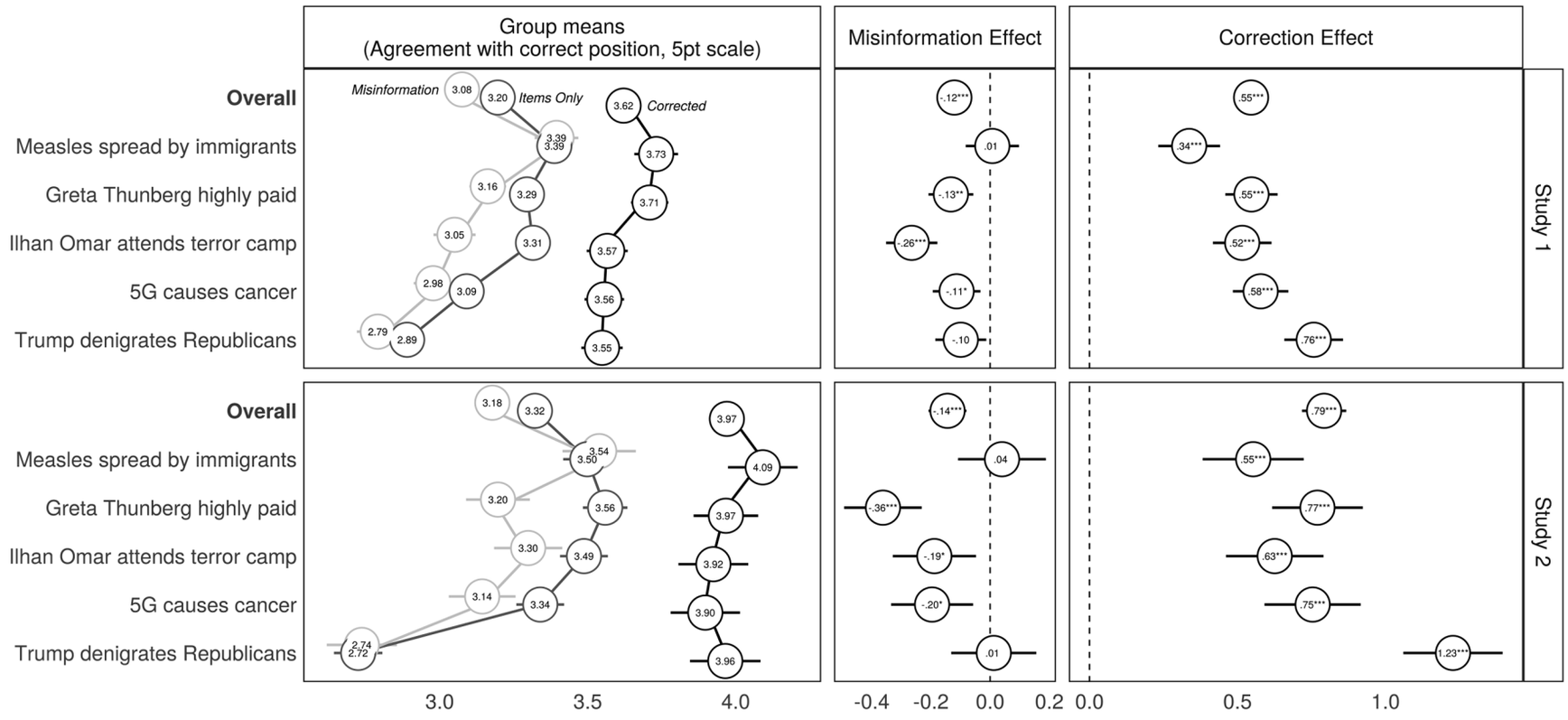

Figure 2. Experimental results with associated 95% confidence intervals. *Left*, group mean values by condition and treatment, with experimental conditions depicted by color. *Middle*, misinformation effects (subtracting the items-only condition from the misinformation condition). *Right*, correction effects (subtracting the misinformation condition from the correction condition). Rows report estimates separately by study. Points summarize the regression models provided in table 1. *** $p < .001$, ** $p < .01$, * $p < .05$.

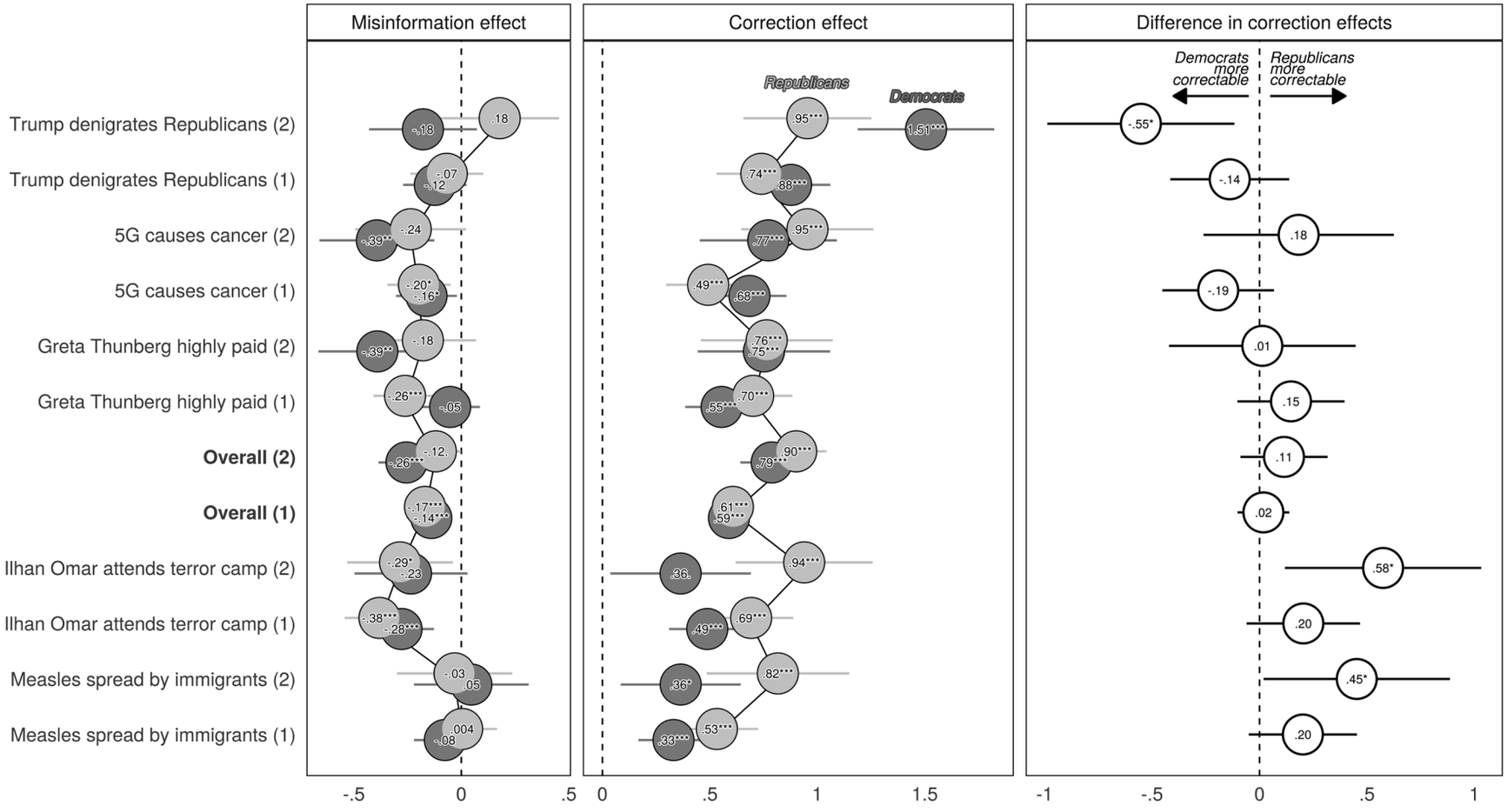

Figure 3. Misinformation effects, correction effects, differences in correction effects, and their 95% confidence intervals, by issue and respondent partisanship. *Y*-axis lists the different experimental issues, with the study number in parentheses. *Left*, misinformation effects (differencing the misinformation and control conditions). *Middle*, correction effects (comparing the misinformation and correction conditions). In the first two panels, light-shaded points report conditional differences among Republicans, dark-shaded points, conditional differences among Democrats. *Right*, difference in correction effects by partisanship. Effects are drawn from the linear models reported in table 10.

by becoming more accurate. In addition, as shown in figure 6, we do not find that the effects of corrections vary with the number of corrections seen.

While overall effects were consistent across parties, corrections to two items that mentioned partisan figures generated larger correction effects among supporters of the out-party. These results, coupled with those regarding immigrants and measles, suggest that partisan differences in effects may be explained by different pretreatment beliefs. We caution, however, that this finding is speculative, with more research needed. Our study is not without its limitations. While our simulated Facebook environment closely resembles Facebook, it is not a comprehensive replication. Conspicuously absent, for example, is the appearance of real-world friends and connections. Genuine social ties may have led to different results. The same might also have been true of more compelling placebo content. For now, however, our results suggest that social media companies, policy makers, and scholars need not resign themselves to the spread of misinformation on social media but can use corrections to rebut it.

ACKNOWLEDGMENTS

We thank Avaaz's Nathan Miller and Fadi Quran, as well as Josh Becker and Alex Coppock for comments. All mistakes are our own.

REFERENCES

Allcott, Hunt, and Matthew Gentzkow. 2017. "Social Media and Fake News in the 2016 Election." *Journal of Economic Perspectives* 31 (2): 211–36.

Bode, Leticia. 2016. "Political News in the News Feed: Learning Politics from Social Media." *Mass Communication and Society* 19 (1): 24–48.

Bode, Leticia, and Emily K. Vraga. 2015. "In Related News, That Was Wrong: The Correction of Misinformation through Related Stories Functionality in Social Media." *Journal of Communication* 65 (4): 619–38.

Brashier, Nadia M., and Elizabeth J. Marsh. 2020. "Judging Truth." *Annual Review of Psychology* 71 (1): 499–515.

Delli Carpini, Michael X., and Scott Keeter. 1996. *What Americans Know about Politics and Why It Matters*. New Haven, CT: Yale University Press.

Einstein, Katherine, and Jennifer Hochschild. 2015. *Do Facts Matter? Information and Misinformation in American Politics*. Norman: University of Oklahoma Press.

Guess, Andrew, Jonathan Nagler, and Joshua Tucker. 2019. "Less than You Think: Prevalence and Predictors of Fake News Dissemination on Facebook." *Science Advances* 5 (1).

Guess, Andrew M., Brendan Nyhan, and Jason Reifler. 2020. "Exposure to Untrustworthy Websites in the 2016 US Election." *Nature Human Behaviour* 4 (5): 472–80.

Hameleers, Michael, and Toni G. L. A. van der Meer. 2020. "Misinformation and Polarization in a High-Choice Media Environment: How Effective Are Political Fact-Checkers?" *Communication Research* 47 (2): 227–50.

Jamieson, Kathleen H. 2018. *Cyberwar: How Russian Hackers and Trolls Helped Elect a President: What We Don't, Can't, and Do Know*. New York: Oxford University Press.

Kalmoe, Nathan P., Raymond J. Pingree, Brian Watson, Mingxiao Sui, Joshua Darr, and Kathleen Searles. 2018. "Crime News Effects and Democratic Accountability: Experimental Evidence from Repeated Exposure in a Multiweek Online Panel." *International Journal of Public Opinion Research* 31 (3): 506–27.

Lodge, Milton, and Charles S. Taber. 2013. *The Rationalizing Voter*. Cambridge: Cambridge University Press.

Nyhan, Brendan, and Jason Reifler. 2010. "When Corrections Fail: The Persistence of Political Misperceptions." *Political Behavior* 32 (2): 303–30.

Pennycook, Gordon, Adam Bear, Evan T. Collins, and David G. Rand. 2020. "The Implied Truth Effect: Attaching Warnings to a Subset of Fake News Headlines Increases Perceived Accuracy of Headlines without Warnings." *Management Science* 66 (11): 4958–79.

Pennycook, Gordon, and David Rand. 2019. "Lazy, Not Biased: Susceptibility to Partisan Fake News Is Better Explained by Lack of Reasoning than by Motivated Reasoning." *Cognition* 188:39–50.

Prior, Markus. 2007. *Post-broadcast Democracy: How Media Choice Increases Inequality in Political Involvement and Polarizes Elections*. New York: Cambridge University Press.

Ryan, Timothy, and Amanda Aziz. 2021. "Is the Political Right More Credulous? Experimental Evidence against Asymmetric Motivations to Believe False Political Information." *Journal of Politics* 83 (3): 1168–72.

Shearer, Elisa, and Katerina Eva Matsa. 2018. "News Use across Social Media Platforms, 2018." Pew Research Center, September 10.

Stroud, Natalie. 2011. *Niche News: The Politics of News Choice*. New York: Oxford University Press.

Swire-Thompson, Briony, Joseph DeGutis, and David Lazer. 2020. "Searching for the Backfire Effect: Measurement and Design Considerations." *Journal of Applied Research in Memory and Cognition* 9 (3): 286–99.

Vraga, Emily K., Leticia Bode, Anne-Bennett Smithson, and Sonya Troller-Renfree. 2016. "Blurred Lines: Defining Social, News, and Political Posts on Facebook." *Journal of Information Technology and Politics* 13 (3): 272–94.

Walter, Nathan, Jonathan Cohen, R. Lance Holbert, and Yasmin Morag. 2020. "Fact-Checking: A Meta-analysis of What Works and for Whom." *Political Communication* 37 (3): 350–75.

Wood, Thomas J., and Ethan Porter. 2019. "The Elusive Backfire Effect: Mass Attitudes' Steadfast Factual Adherence." *Political Behavior* 41 (1): 135–63.