Unchecked vs. Uncheckable: How Opinion-Based Claims Can Impede Corrections of Misinformation

Nathan Walter & Nikita A. Salovich

View supplementary material

Published online: 04 Feb 2021.

Submit your article to this journal

View related articles

View Crossmark data

Routledge
Taylor & Francis Group

Check for updates

Unchecked vs. Uncheckable: How Opinion-Based Claims Can Impede Corrections of Misinformation

Nathan Walter and Nikita A. Salovich

Department of Psychology, Northwestern University, Chicago, IL, USA

ABSTRACT
Although the prominence of fact-checking in political journalism has grown dramatically in recent years, empirical investigations regarding the effectiveness of fact-checking in correcting misperceptions have yielded mixed results. One understudied factor that likely influences the success of fact-checking initiatives is the presence of opinion statements in fact-checked messages. Recent work suggests that people may have difficulty differentiating opinion- from fact-based claims, especially when they are congruent with pre-existing beliefs. In three experiments, we investigated the consequences of opinion-based claims to the efficacy of fact-checking in correcting misinformation regarding gun policy. Study 1 ($N = 152$) demonstrated that fact-checking is less effective when it attempts to correct statements that include both fact- and opinion-based claims. Study 2 ($N = 561$) replicated and expanded these findings showing that correction is contingent on people's ability to accurately distinguish facts from opinions. Study 3 ($N = 389$) illustrated that the observed effects are governed by motivated reasoning rather than actual inability to ascertain fact-based claims. Together these results suggest that distinguishing facts from opinions is a major hurdle to effective fact-checking.

Recent calls have urged researchers to shift their focus from a research agenda that tries to conceptualize and typologize *fake news* to studies that outline and delineate the boundary conditions of influence exerted by misinformation and its correction (Weeks & Gil de Zúñiga, 2021). The current inquiry heeds this recommendation by outlining the opportunities and limitations of fact-checking—one of the most popular and rapidly developing methods used by

political journalists and other gatekeepers to correct misinformation and infuse public discourse with facts (Nieminen & Rapeli, 2019). Although the industry of fact-checking has grown enormously to the point that some view it as a staple of U.S. elite journalism (Graves, 2018), the empirical evidence regarding the benefits of fact-checking in supporting correct political knowledge and beliefs are far from being encouraging (Walter et al., 2020). Ostensibly, the exclusive focus on fact-based statements and the nonpartisan nature that characterizes fact-checkers should have maximized their efficacy to correct misperceptions while minimizing partisan bias. With that in mind, however, studies find that across a variety of topics fact-checking messages are interpreted through a partisan lens, with the public being unwilling to accept factual information that runs counter to their views (e.g., Jarman, 2016; Thorson, 2016). Given this discrepancy, there is a pertinent need to go beyond such previous studies and understand why fact-checking, the gold-standard of political journalism, fails far too often at correcting misperceptions.

Informed by motivated reasoning (Kunda, 1990; Taber & Lodge, 2006) and the implied truth effect (Pennycook et al., 2020), the current paper includes three experiments that test whether the effectiveness of fact-checking in correcting misinformation regarding gun policy is contingent on people's ability to discern fact- from opinion-based statements. Specifically, Study 1 investigates whether the integration of opinion-based statements within a false claim can weaken the effects of counter-attitudinal fact-checking. Study 2 expands these results by testing whether the effectiveness of fact-checking is contingent on the ability to distinguish between fact- and opinion-based statements. Study 3 complements the findings by assessing whether the confusion between fact- and opinion-based statements can be traced back to motivated reasoning or an actual inability to identify verifiable claims. Each experiment incrementally expands on the previous to provide a theory-driven explanation for why fact-checking often fails and what can be done to address it.

On political misinformation and its remedies

Although political misinformation is as old as democracy itself (Grant, 1995), the deteriorating control of traditional gatekeepers coupled with the anonymity, immediacy, polarization, and financial model of the current media ecology provide fertile ground for the proliferation of false messages (Allcott & Gentzkow, 2017). From global geopolitical events such as the Brexit referendum (Bennett & Livingston, 2018) and the 2016 U.S. presidential elections (Allcott & Gentzkow, 2017) to minor political controversies regarding the relative sizes of crowds at presidential inaugurations (Schaffner & Luks, 2018), the diffusion of misinformation poses substantial threats to a well-functioning democracy (Jamieson, 2018).

Given the sophistication of misinformation and the challenges to root out its consequences, it is unsurprising that journalists and other gatekeepers have been developing approaches to mitigate and contain the spread of falsehoods (Margolin et al., 2018). One of the more recent innovations designed to address counterfactual political information has been fact-checking. The last few years have seen a global surge in political fact-checking. Defined as "the practice of systematically publishing assessments of the validity of claims made by public officials and institutions with an explicit attempt to identify whether a claim is factual" (Walter et al., 2020, p. 351), fact-checking organizations such as FactCheck.org and PolitiFact are routinely embraced by news organizations and political elites (Graves, 2018).

The turn toward fact-checking in recent years is not indicative, however, of its success in reducing the perceived accuracy of false claims or tackling the spread of political misinformation (Egelhofer & Lecheler, 2019). Indeed, the current state of empirical evidence regarding the effects of fact-checking on political beliefs lends itself to a variety of interpretations, from steadfast adherence to fact-checked information (e.g., Wood & Porter, 2019), to null findings (e.g., Garrett & Weeks, 2013) or persistence of distorted attitudes in the face of correction (e.g., Thorson, 2016). The heterogeneity of empirical evidence is also echoed in a recent meta-analysis that summarized findings from 30 studies where fact-checking was used to correct political misinformation (Walter et al., 2020). The synthesized findings suggest that effects are highly contingent (d = 0.17– 0.52) on a variety of factors, including participants' political ideology and the structure of false information. This observation seems to align with a growing urgency to shift the focus of political communication from bottom-line effects to trying to understand *when* messages matter (Weeks & Gil de Zúñiga, 2021). Informed by the literature of motivated reasoning and the implied truth effect, the following sections identify two potential explanations for the (in) effectiveness of fact-checking in correcting people's beliefs.

Fact-checking and motivated reasoning

Motivated reasoning theory originated in the basic premise that people are guided by different needs when they encounter, process, and make decisions based on information (Maslow, 1954). Studies have differentiated between directional goals (i.e., attempting to be consistent with a preexisting belief) and accuracy goals (i.e., attempting to be correct and precise) (Chaiken et al., 1989; Kunda, 1990). In the political context, accuracy goals appear to be subordinate to individuals' predispositions to reach conclusions in a particular direction (Kunda, 1990). As such, counter-attitudinal information is often avoided, derogated, or rejected, whereas information that aligns with preexisting beliefs is judged as probative and accepted without considerable scrutiny (Taber & Lodge,

2006). Along with the emphasis on distinct motivations, the theory suggests that biased processing also depends on people's involvement with the issue (Kunda, 1990), as well as the perceived credibility of the source of information (Kuru et al., 2017). Simply put, individuals are likely to be especially vulnerable to motivated reasoning when encountering information from outgroup members that challenges their value-laden beliefs (Druckman & McGrath, 2019). For instance, objective information that attempts to correct false beliefs associated with gun laws is likely to be trusted if it aligns with one's ideology and discounted when it contradicts it.

Ostensibly, the tendency to rely on directional goals when processing value-laden information should be less pronounced in the context of fact-checking. Unlike other journalistic strategies employed to correct political misinformation, including retractions and ad watches, fact-checkers are nonpartisan, nonprofit, independent organizations that are largely unaffiliated with mainstream media or partisan sources (FactCheck.org, n.d.). Yet, when it comes to partisanship, the ability of counter-attitudinal fact-checking to challenge beliefs is remarkably consistent with findings from any other form of corrective communication. For instance, Thorson (2016) demonstrated that partisans are unwilling to accept fact-checking messages that run counter to their view, even when misinformation is corrected immediately. A similar conclusion was reached by Jarman (2016), indicating that the fact-checking message was subject to partisan interpretation with corrective information used to actually *solidify* the original false opinion. As summarized by Amazeen et al. (2018), if "a tertiary goal of fact-checking is to have positive effects on journalism, it is noteworthy—yet concerning—that people feel more favorably toward fact-checkers when they correct the opposition and less so when they correct one's own party" (p. 42). Hence, the limitations of fact-checking in addressing false beliefs can be, in part, explained by motivated reasoning that governs processing of political information. More recently, however, researchers have proposed an additional explanation for the limited influence of fact-checkers (e.g., Merpert et al., 2018). According to this account, the qualities that make fact-checking unique, including its exclusive focus on *fact-based statements*, can also substantially limit its effects and even backfire.

Fact-checking and the implied truth effect

The idea that fact-checking messages are subject to predisposed biases that can solidify (rather than correct) preexisting beliefs warrants further investigation into the characteristics of fact-checking messages. One of the distinct aspects of fact-checkers is their exclusive focus on *checkable* political statements or claims that can be factually verified. Prima facie, the exclusive focus on fact-based statements—something that is capable of being proved or disproved by objective evidence—should elevate the discourse and highlight

the need to support political arguments with evidence and data. At the same time, however, the exclusive focus on fact-based claims can also pose a considerable risk, since a large portion of political talk involves unverifiable statements or opinions outside the realm of fact-checking (Amazeen, 2016). Hence, there is a risk that political entities can pollute the public discourse and pass fact-checking examinations by targeting a "gray area" of statements that sound like fact-based claims but are actually unverifiable opinions.

People's general tendency to confuse factual and opinion statements amplifies this concern. For example, a recent Pew Research Center survey that exposed U.S. adults to five fact-based statements (e.g., "Health care costs per person in the U.S. are the highest in the developed world") and five opinion-based statements (e.g., "Increasing the federal minimum wage to 15 USD an hour is essential for the health of the U.S. economy") found that only 26% were able to accurately classify all statements (Pew Research Center, 2018a). Similarly, when fact- and opinion-based statements were integrated into a fictional political speech, on average, participants were able to correctly identify only 69% of all eight fact-based statements (Merpert et al., 2018).

Setting aside the specific reasons that underlie these gaps, if some statements are flagged as inaccurate and others are not (either because they are accurate or because they are opinions), individuals may incorrectly infer that unflagged messages are trustworthy (Clayton et al., 2020).

Pennycook et al., (2020) coined the term "implied truth effect" to describe a phenomenon whereby the existence of headlines tagged as "false" increases the perceived accuracy of untagged fake headlines. The idea is rooted in Bayesian models of cognition (Tenenbaum, Griffiths, & Kemp, 2006), wherein people's comprehension of and judgments about information reflect inferences from base rates acquired from previous experiences. For example, as a general bias, people are more likely to assume that a new piece of information is true than false, given that most information encountered in daily life is true (Brashier & Marsh, 2020). Bayes' rule also predicts that, in the presence of tagged and untagged false information, people's beliefs in information "tagged" as false would decrease, while beliefs in information with no tag would increase —the latter being the implied truth effect. Indeed, the implied truth effect was empirically supported in a series of five studies showing that the presence of a tag that indicated that a story was false caused untagged stories to be seen as more accurate, irrespective of whether these stories were actually true or not (Pennycook & Rand, 2020).

As further support for the implied truth effect, fact-checking has been shown to have weaker effects on beliefs when not all of the original statements were explicitly flagged as false (Walter et al., 2020). When fact-checkers' verdicts were ambiguous (e.g., "mostly false"), people interpreted the equivocation to benefit their preexisting beliefs. Considering the exclusive focus on

fact-based claims, the absence of a disconfirmation becomes ambiguous: Does it mean that the claims were successfully verified, or does it mean that the information cannot be fact-checked? According to this logic, the implied truth effect presents a challenge to fact-checking because not all statements can be explicitly tagged as false. Given the demonstrated difficulty in differentiating opinion- from fact-based claims, people may erroneously perceive unaddressed opinion statements as "true," or at the very least, reject the idea that all the information present in the original message is false. Either way this would increase the perceived validity of the original message and decrease the efficacy of the fact-check. As summarized by Merpert et al. (2018), for "fact-checking to be successful, the first step is to accurately identify if statements contain facts that can be checked … [w]ithout being able to identify if information presented is fact or opinion, fact-checkers and the public alike would have no basis on which to begin their scrutinizing" (Merpert et al., 2018, p. 49).

Overview of studies

The article presents three studies investigating the opportunities and limitations of fact-checking regarding misinformation associated with gun policy. These studies correspond with three primary questions: (1) What is the effect of motivated reasoning and the implied truth effect on fact-checking? (2) What mechanisms can explain the limitations of fact-checking? and (3) What are the motivations that underlie acceptance and rejection of fact-checking messages? In particular, Study 1 examines the role played by (a) ideological-congruence of misinformation and (b) the checkability (facts vs. opinions) of the content on readers' agreement with the fact-checker's conclusion. Study 2 replicates the results, addresses important limitations, and tests a theoretical mechanism, wherein the efficacy of fact-checking is governed by individuals' ability to distinguish fact- from opinion-based statements. Study 3 expands these findings by examining whether the limited ability to distinguish statements results from the need to be consistent with preexisting beliefs or an actual confusion between facts and opinions.

In order to test these assumptions in a real-world context, all three studies focus on fact-checking of misinformation related to gun policy. Gun-related policy has emerged as one of the most polarizing issues in modern American politics, with Democrats being nearly three times as likely as Republicans to say gun laws in the U.S. should be stricter than they are today (reflecting agreement at 80% and 28%; Pew Research Center, 2018b). Hence, the focus on gun policy will ensure a challenging test for the efficacy of fact-checking. Relatedly, gun policy is well-positioned to serve as the context for its strong association with political parties. Keeping in mind

that the current study focuses on ideological-congruence of misinformation, the focus on guns permits a valid assumption that if someone self-identifies as a Democrat they are likely to support stricter gun laws compared with their Republican counterparts. As argued by Silver (2012), "[w] hether someone owns a gun is a more powerful predictor of a person's political party than her gender, whether she identifies as gay or lesbian, whether she is Hispanic, whether she lives in the South or a number of other demographic characteristics" (Silver, 2012). Finally, the focus on gun policy is also justified on the grounds of external validity, as statements about guns represent one of the most fact-checked areas in political discourse. For instance, according to aggregate scores, PolitiFact assessed the veracity of nearly 600 individual statements related to gun policy, finding that only 83 (14%) were completely true (PolitiFact, 2019).

Study 1: Fact-checking in the face of ideological-congruence and opinion statements

The first study examined the postulated role played by ideological-congruence and opinion statements in receptivity to fact-checking. Based on previous research on motivated reasoning and the implied truth effect, two hypotheses and one research question were posed:

H1: The ideological-congruence of misinformation reduces the effectiveness of subsequent fact-checking on (a) message-accuracy and (b) speaker's credibility, such that weaker effects are observed for participants who are exposed to counter-attitudinal fact-checking compared to pro-attitudinal fact-checking.

H2: Message-checkability influences the effectiveness of fact-checking, such that exposure to false claims that include both fact- and opinion-based statements results in weaker effects on (a) message-accuracy and (b) speaker's credibility, compared to false statements that include only fact-based claims.

RQ1: Is there an interaction effect between ideological-congruence of misinformation and message-checkability on (a) message-accuracy and (b) speaker's credibility?

Method[1]

Design and participants

The study employed a 2 (misinformation-congruence: pro-attitudinal or counter-attitudinal) X 2 (message-checkability: combination of fact- and opinion-based statements or only fact-based statements) factorial design. The goal was to recruit 150 participants based on a priori power analysis (G*Power 3; Faul et al., 2007), suggesting that this sample size is needed to detect an effect size comparable with a recent meta-analysis of fact-checking (d = 0.40, α = .05, 1-β = .80; Walter et al., 2020). One hundred and sixty-eight consenting participants were recruited from a Qualtrics panel. Later, 16 participants were removed due to an unobtrusive measure of time elapsed that indicated that they had not read the study material, bringing the effective sample size to N = 152. The average age of participants was 48.91 (SD = 15.33) and 75% were female. Participants indicated that, on average, they completed 13.91 (SD = 3.52) years of schooling, and 129 (84.9%) identified as white, followed by Black (11, 7.2%), Hispanic (4, 2.6%), American Native (2, 1.3%), and Asian (3, 2.0%). Eighty participants self-identified as Democrats and 72 identified as Republicans, with ideology ranging from *strong liberal* (1) to *strong conservative* (7) (M = 4.11, SD = 2.05).

Procedure and material

To assess misinformation-congruence, all consenting participants were required to indicate their political affiliation (Democrat, Republican, Independent, or Other). Then, only those who self-identified as Democrats or Republicans were randomly assigned to receive either an article in favor of gun rights (advocating for less restrictions on gun ownership) or an article in favor of gun control (advocating for more restrictions on gun ownership) (312–443 words depending message-checkability; see Appendix A).[2] Adapted from Graves et al., 2018), the news article, attributed to a "Pennsylvania-based newspaper," covered the "London Global Forum on Guns," focusing on an

---

[1]All three studies were approved by the Institutional Review Board at Northwestern University (Study number: STU00209397).

[2]Given concerns over a potential pretesting effect, we decided to use participants' political affiliation, rather than directly measuring participants support for gun-related policy, to compute the ideological-congruence variable. To ensure that political affiliation can accurately predict participants' views on gun-related policy, the final part of the questionnaire directly measured participants' views on guns. Specifically, participants were asked to "Choose the option that best depicts how [they] feel about gun laws," ranging from (1) "Gun laws should be more restrictive (harder to get and hold guns)" to (7) "Gun laws should be less restrictive (easier to get and hold guns)."The results of an independent samples t-test indicated a strong effect of political affiliation on support for gun-related policy (t(150) = 5.58, p < .001, d = 0.86), with Republicans much more likely to support gun rights (M = 5.37, SD = 2.82), compared to Democrats (M = 3.00, SD = 2.68).

interview with the founder of a fictitious group – the Responsible Firearms International (RFI) group (in the gun rights condition) or the Citizens for Gun Free Streets (CGFS) group (in the gun control condition). Using the interview format, both versions of the article included five fact-based statements (e.g., "virtually 100% of shooters have some kind of a mental illness history . . ."). Depending on the experimental condition, half of the participants read the same article that also included five opinion-based statements (e.g., "we don't have enough laws that restrict guns in America"). Following a classification used in Merpert et al. (2018), fact-based claims were identified as those that relate to historical data, comparisons, legality, and statistics, whereas opinion statements were those pertaining to general future projections, values, and personal judgments. Notably, both versions of the article were made to mirror each other as closely as possible, by presenting logically equivalent false statements. For example, the gun rights version included the fact-based claim that "virtually 100% of shooters have some kind of a mental illness history," whereas the gun control version included in the claim that "only 33% of shooters have some kind of a mental illness history." In truth, 68% of gun crime involves shooters with a record of mental illness. Hence, all five fact-based statements across both articles were equivalently false.

The decision to focus on fact-checking that challenges false statements as opposed to fact-checking that confirms accurate statements was guided by three considerations. First, from an external validity perspective, approximately 86% of fact-checked statements regarding gun policy yield negative verdicts (PolitiFact, 2019). Second, based on previous studies (e.g., Bode & Vraga, 2015), we expected to find a stronger effect of fact-checking for debunking messages compared to confirming messages. Third, the focus on false statements allows for a clearer comparison of the results with extant literature, since only a limited number of studies evaluated the impact of favorable fact-checking (Walter et al., 2020).

After reading the news article, participants completed an unrelated distractor task for approximately five minutes (20 trivia questions unrelated to guns). Beyond consistency with previous studies (e.g., Ecker et al., 2015), the inclusion of a distractor task resembles situations in the real world, as time may pass between getting information and then learning that it was false (Thorson, 2016). Following the distractor task, all participants were exposed to a fact-checking message (714 words; see Appendix B), ostensibly from PolitiFact, which directly contradicted the five fact-based statements made in the original article, rating all statements as "false" (both in-text and using a "Truth-O-Meter"— a thermometer-like visual representation of how false an article is). We chose to attribute the corrective information to PolitiFact, as it was previously identified as the most commonly utilized fact-checker in extant research (Walter et al., 2020). Keeping in mind that the false statements were designed to mirror each other, we were able to use

the same fact-checking information to contradict the false statements in both versions. This procedure concluded with a battery of questions designed to measure the manipulation check and the research outcomes. On average it took each participant 13 minutes (SD = 14.7 min) to complete the study.

Measures

A *manipulation check* measured whether participants accurately identified the speaker's position on guns in the original statement they read. Specifically, participants were asked whether the position of the speaker (Erica Taylor in both versions) reflects gun-rights, gun-control, or both perspectives. The *Speaker's credibility* measure was adapted from Flanagin and Metzger (2000), with participants being instructed to respond to the prompt "Thinking about the news article that you just read. Overall, you would say that Erica Taylor (the speaker at the "Global Forum on Guns) is:" The five semantic differential items, ranging from *low credibility* (1) to *high credibility* (7) included "Untrustworthy-Trustworthy" and "Dishonest-Honest" (M = 3.53, SD = 1.92, α = .96). Finally, *Message accuracy* was gauged with a single-item ("To the best of your knowledge, how accurate are the claims made by Erica Taylor [the speaker at the Global Forum on Guns]?"). Response options ranged from *completely inaccurate* (1) to *completely accurate* (7) (M = 3.45, SD = 1.92).

Results and discussion

The message manipulation proved successful, as those in the gun rights condition were more likely to identify the stimulus as reflecting favorable views of gun rights (52, 70.3%; adjusted residual = 7.0), compared to gun control (13, 17.6%; adjusted residual = −6.9), or equal representation for both perspectives (9, 12.2%; adjusted residual = −0.10), whereas those in the gun control message condition viewed the stimulus as advocating for gun control (57, 73.1%; adjusted residual = 6.9), compared to gun rights (11, 14.1%; adjusted residual = −7.0), or balanced representation (10, 12.8%; adjusted residual = 0.10); χ^2 (2, N = 152) = 54.33, r_c = .60, p < .0005.

The hypotheses and research questions were tested with a Multivariate Analysis of Variance (MANOVA), treating misinformation-congruence and message-checkability as fixed factors and message accuracy and speaker's credibility as outcomes. The MANOVA indicated a significant omnibus effect for misinformation-congruence (Wilk's λ = .85, F (2,147) = 12.73, p < .0005, η_p^2 = .15). Examining its univariate effects, the model recorded a significant influence on message accuracy (F (1,152) = 18.46, p < .001, η_p^2 = .11) and speaker's credibility (F (1,152) = 25.14, p < .001, η_p^2 = .15). As indicated in Figure 1 and in support for H1, when an ideologically congruent false
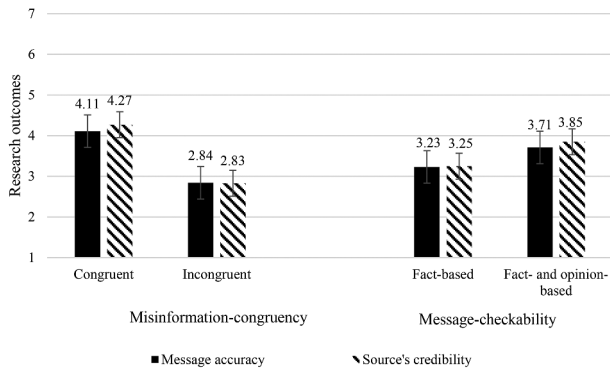
Figure 1. Means (and 95% confidence intervals) for research outcomes by experimental condition (study 1).

statement was debunked by a fact-checker, participants were less likely to view the false message as inaccurate and challenge the credibility of its source.

Further, the MANOVA also revealed a significant main effect of message-checkability (Wilk's λ = .97, F (2,147) = 3.73, p = .034, η_p^2 = .04). An examination of the univariate effects suggested that message-checkability exerted a significant influence both on message accuracy (F (1,152) = 3.42, p = .04, η_p^2 = .02) and on source's credibility (F (1,152) = 4.33, p = .03, η_p^2 = .03). Figure 1 illustrates the pattern of results showing that exposure to misleading information that included both fact- and opinion-based statements substantially reduced the effectiveness of a subsequent fact-checking, compared to exposure to a misleading message the included only fact-based statements. Thus, H2 was supported.

With regard to RQ1, there was no indication of an interaction between misinformation-congruence and message-checkability (Wilk's λ = .99, F (2,147) = 0.12, p = .89, η_p^2 = .00). Likewise, the univariate analysis did not identify a significant interaction effect on message accuracy (F (1,152) = 0.13, p = .72, η_p^2 = .00) or speaker's credibility (F (1,152) = 0.24, p = .63, η_p^2 = .00).

In sum, fact-checking appears less effective when the target message contains opinion-based statements in addition to the facts they are attempting to correct. This effect persists regardless of the message's congruence with the reader's partisan beliefs.

Study 2: Message checkability detection and fact-checking

After showing the expected influence of misinformation-congruence and opinion-based statements on perceived accuracy and speaker's credibility,

the aim of Study 2 was to propose a theoretical mechanism and address limitations on several fronts. First, Study 1 used a convenience sample skewed in terms of gender and race/ethnicity, whereas Study 2 utilized a more balanced sample. Second, while Study 1 exposed all participants to misinformation that was later fact-checked, Study 2 randomly assigned participants either to receive a fact-checking message or not, allowing assessment of the influence of exposure to fact-checking. Third, to ensure that participants accept the notion of fact-checkers as nonpartisan organizations, Study 2 included a relevant manipulation check. Fourth, Study 2 eliminated the confounding threats in Study 1 by keeping message length equal across all conditions. Fifth, whereas Study 1 assumed that it will be more difficult to correct misinformation that includes both fact- and opinion-based statements, Study 2 directly measured participants' ability to discern the two types of statements as a possible moderator. These modifications and the preceding literature review lead to the following hypotheses:

Hypotheses

H3: Exposure to ideologically congruent misinformation positively affects judgments of (a) message-accuracy and (b) speaker's credibility.

H4: Exposure to fact-checking negatively affects judgments of (a) message-accuracy and (b) speaker's credibility.

H5: There is an interaction between misinformation-congruence and exposure to fact-checking, such that fact-checking will be more effective when attempting to correct counter-attitudinal misinformation, compared with pro-attitudinal misinformation.

H6: The effectiveness of fact-checking is contingent on message-checkability detection, such that exposure to fact-checking exerts more influence when individuals have greater ability to discern fact- from opinion-based statements.

Method

Design and participants

Study 2 employed a 2 (misinformation-congruence: pro-attitudinal or counter-attitudinal) x 2 (fact-checking: present or absent) x 2 (message-checkability detection: with or without[3]) between-subjects factorial design. Broadly speaking, the experimental procedure included five distinct stages: (a) exposure to pro- or counter-attitudinal misinformation; (b) distractor task; (c) exposure to a message-checkability detection task; (d) exposure to fact-checking that debunks the original statement; and (e) measurement of research outcomes.

Participants were 561 (284, 50.6% females) U.S. residents recruited from Qualtrics Panels.[4] Of them, 294 (52.4%) identified as Democrats and 267 (47.6%) identified as Republicans. Their age ranged from 18 to 94 years with a mean of 45.97 (SD = 16.97). We utilized the 2018 census to match quotas to represent the national population in terms of race/ethnicity, education, and geography. The majority of the sample were White (342, 61.0%), followed by Hispanic and Latino Americans (101, 18.0%), Black or African American (76, 13.5%), Asian (32, 5.7%), Native Americans and Alaska Natives (3, 0.5%), and the mean years of schooling were 13.29 (SD = 2.01).

Procedure and material

The procedure closely followed Study 1 with several notable differences. First, both the gun rights and gun control versions of the misinformation included five fact-based and five opinion-based statements. Second, following the distractor task, participants were randomly assigned either to receive a message-checkability detection task or not receive this task at all.[5] Considering the possibility that such measurement is likely to be influenced by the perceived accuracy of false statements, we decided to

---

[3]Half of the participants in the condition that included a message-checkability detection task received feedback on their performance and half did not receive feedback. Considering the fact that performance feedback did not have direct or moderated effects on any of the research outcomes, including on message accuracy (b = −.59, SE = .56, p = .29, 95% CI [−1.70, .51], ΔR^2 = .01) and speaker's credibility (b = −.72, SE = .56, p = .20, 95% CI [−1.82, .38], ΔR^2 = .01), the two conditions were combined.

[4]Correcting for the increase in the number of experimental conditions, the same power analysis was utilized to calculate the required sample size for Study 2 (d = 0.40, a = .05, 1-β = .80) resulting in N = 520.

[5]This was done to examine a potential confound, whereby the mere inclusion of the task might independently influence the research outcomes. As the MANOVA with message-checkability detection task (with/without) indicated, inclusion of the task did not significantly influence the research outcomes (Wilk's λ = .99, F (3, 826) = 1.19, p = .31, η_p^2 = .004).

introduce it prior to exposing participants to the fact-checking message. Likewise, this procedure more closely corresponds with how people process information in the real world, as judgments that separate fact-based from opinion-based claims tend to occur immediately after information processing, as opposed to a post-hoc decision that follows fact-checking. Adapted from previous work (e.g., Merpert et al., 2018; Pew Research Center, 2018a), the message-checkability detection task included ten randomized rounds, with each round presenting one sentence from the original news article and asking participants to decide whether it represents a factual statement ("a claim that can be objectively verified") or an opinion statement ("a claim that cannot be verified").[6] Prior to this study, the checkability detection task was successfully piloted (N = 32) on students from a large Midwestern U.S. university. The purpose of the pilot was to identify statements that are significantly easier or more difficult to classify. The pilot did not surface any problems, and thus we proceeded to include all 10 statements. After performing the message-checkability detection task, only half of the participants were randomly assigned to view a fact-checking message.

Measures

The manipulation check from Study 1 was supplemented with an additional measure that assessed people's acceptance of PolitiFact as a nonpartisan source. Those who were exposed to a fact-checking correction were asked to identify whether PolitiFact is a "non-partisan," "liberal," or "conservative fact-checking organization" (Amazeen et al., 2018). Beyond *speaker's credibility* (M = 4.39, SD = 1.78, α = .93), and message accuracy (M = 4.22, SD = 1.84), *message-checkability detection* was gauged by summing up the number of correct responses on the task. The scores ranged from 0 (low ability) to 10 (high ability) with a mean of 6.26 (SD = 1.71) statements correctly classified as fact- or opinion-based.

Results and discussion

The message manipulation was successful. Participants in the gun rights condition were more likely to identify the stimulus as reflecting favorable

---

[6]The specific instructions were: "The original article published in The News-Item included an interview with the political activist, Erica Taylor. The interview included both factual statements that can be objectively verified, as well as opinion-statements that cannot be verified. The following questions test your ability to classify statements as either factual or opinion. Regardless of how knowledgeable you are about each topic, would you consider each statement to be a factual statement (whether you think it is accurate or not) or an opinion statement (whether you agree with it or not)?".

views on gun rights (259, 62.6%; adjusted residual = 14.7), compared with gun control (98, 23.7%; adjusted residual = −14.9), or an equal representation for both perspectives (57, 13.8%; adjusted residual = 0.90); whereas those in the gun control message condition tended view the stimulus as advocating for gun control (313, 75.2%; adjusted residual = 14.9), compared with gun rights (54, 13%; adjusted residual = −14.7), or a balanced representation (49, 11.8%; adjusted residual = −0.90); χ^2 (2, N = 830) = 247.34, r_c = .55, p < .0005. Similarly, the manipulation of exposure to a nonpartisan fact-checking message was successful, as indicated by participants' agreement that PolitiFact is a nonpartisan organization (262, 64.2%; adjusted residual = 126.0), as opposed to a liberal (105, 25.7%; adjusted residual = −31.0), or a conservative organization (41, 10%; adjusted residual = −95.0); χ^2 (2, N = 408) = 190.16, p < .0005.

After ensuring that the manipulations were successful, a MANOVA revealed a significant and substantial main effect for misinformation-congruence (Wilk's λ = .94, F (2, 556) = 16.68, p < .0005, η_p^2 = .06), as well as a main effect for fact-checking (Wilk's λ = .95, F (2, 556) = 13.61, p < .0005, η_p^2 = .05), and a nonsignificant multivariate interaction between misinformation-congruence and fact-checking (Wilk's λ = .99, F (2, 556) = 1.93, p = .15, η_p^2 = .01). Specifically, as predicted by H3, exposure to a pro-attitudinal message resulted in greater perceived accuracy (M = 4.72, SD = 1.70) and speaker's credibility (M = 4.92, SD = 1.64), compared to exposure to a counter-attitudinal message (M = 4.04, SD = 1.98; M = 4.09, SD = 1.91; M = 3.33, SD = 1.94; respectively). Under H4, exposure to fact-checking significantly reduced the perceived accuracy of misinformation (M = 3.99, SD = 1.99 vs. M = 4.79, SD = 1.66) and its source's credibility (M = 4.21, SD = 1.88 vs. M = 4.81, SD = 1.71).

In contrast to H5, there was no significant interaction between misinformation-congruence and exposure to fact-checking for speaker's credibility (F (1, 561) = 2.55, p = .11, η_p^2 = .01). Although the omnibus interaction effect was nonsignificant, the results of the post-hoc test in Table 1 revealed a more pronounced decrease in perceived credibility for those reading ideologically congruent fact-checking compared to

Table 1. Means (+SDs) for research outcomes by experimental condition (Study 2).

	Experimental condition			
	Congruent misinformation with fact-checking	Congruent misinformation without fact-checking	Incongruent misinformation with fact-checking	Incongruent misinformation without fact-checking
Speaker's credibility	4.52 (1.66)[a]	5.36 (1.49)[b]	3.91 (2.06)[c]	4.27 (1.75)[ac]
Message accuracy	4.20 (1.77)[a]	5.29 (1.42)[b]	3.79 (2.17)[ac]	4.29 (1.73)[ac]

Means with differing scripts within outcome variables are significantly different at the p < .05 level with Tukey HSD post-hoc test.

ideologically incongruent fact-checking (M_{diff} = 0.61, SE = .21, p = .02, 95% CI [.08, 1.15]). Similarly, there was no significant interaction between misinformation-congruence and exposure to fact-checking on message accuracy (F (1, 561) = 3.81, p = .05, η_p^2 = .01). Notably, while participants who were exposed to ideologically congruent fact-checking reported on a slightly lower level of perceived message accuracy (M = 3.79, SD = 2.17) compared to those who read an ideologically incongruent fact-checker (M = 4.20, SD = 1.77), those differences were not statistically significant (M_{diff} = 0.41, SE = .21, p = .22, 95% CI [−.14, .96]).

To test the possibility that counter-attitudinal fact-checking is contingent on participants' ability to distinguish fact- from opinion-based statements, we conducted a moderation analysis that focused only on those participants that were exposed to counter-attitudinal fact-checking, using PROCESS (Hayes, 2018; Model 1). The analysis treated exposure to fact-checking (yes vs. no) as the focal predictor, message-checkability detection score was entered as a continuous moderator, and perceived accuracy and source's credibility were entered as outcomes. In line with our prediction, the ability to discern fact- from opinion-based statements significantly moderated the effects of counter-attitudinal fact-checking; b = −.35, SE = .14, p = .01, 95% CI [−.61, −.08], ΔR^2 = .03.

Probing this moderation with the Johnson-Neyman technique revealed a pattern in which exposure to counter-attitudinal fact-checking significantly reduced the perceived accuracy of false statements only for those who were able to correctly classify at least 7.86 statements as facts or opinions (only
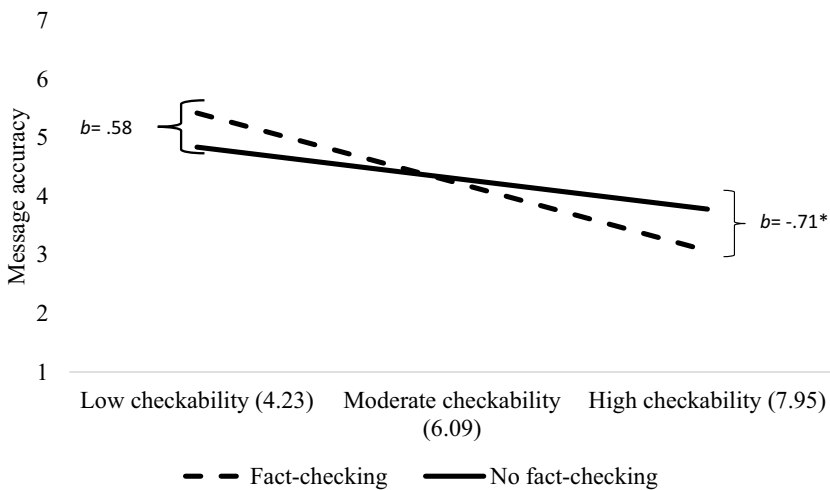

Figure 2. Estimated means of perceived accuracy at different levels of checkability detection (study 2).

23.53% of those exposed to counter-attitudinal fact-checking) (Figure 2). By the same token, the regions of significance also suggested that fact-checking actually backfired among participants who were able to accurately classify fewer than 3.48 (10.16% of those exposed to counter-attitudinal fact-checking). For those participants, exposure to fact-checking had the unintended effect of increasing the perceived accuracy of false information.

A similar analysis with speaker's credibility as the outcome also recorded a significant moderation for message-checkability detection; $b = -.43$, $SE = .14$, $p = .002$, 95% CI [−.71, −.16], $\Delta R^2 = .04$. Again, inspection of the significance areas with the Johnson-Neyman technique identified that the effect of counter-attitudinal fact-checking on perceived credibility was negative and significant, albeit only for those who were able to successfully classify fact- and opinion-based statements (>8.15, 11.75% of those exposed to counter-attitudinal fact-checking) (Figure 3). See Table 1 for a complete outline of the direct and moderated effects of counter-attitudinal fact-checking. The same analysis also revealed a troubling pattern among participants who were not successful in classifying these statements (<5.06, 38.5% of those exposed to counter-attitudinal fact-checking). Specifically, for those participants, a fact-checking message designed to debunk false claims actually increased the perceived credibility of the source responsible for the falsehoods. Finally, in the case of pro-attitudinal fact-checking, an equivalent PROCESS model did not reveal a significant moderation of message-checkability detection on message accuracy ($b = -.04$, $SE = .15$, $p = .787$, 95% CI [−.33, .25], $\Delta R^2 = .00$) and speaker's credibility ($b = -.08$, $SE = .15$, $p = .576$, 95% CI [−.38, .21], $\Delta R^2 = .00$) (see Table 2 for a complete outline of the direct and moderated effects of counter-attitudinal fact-checking).
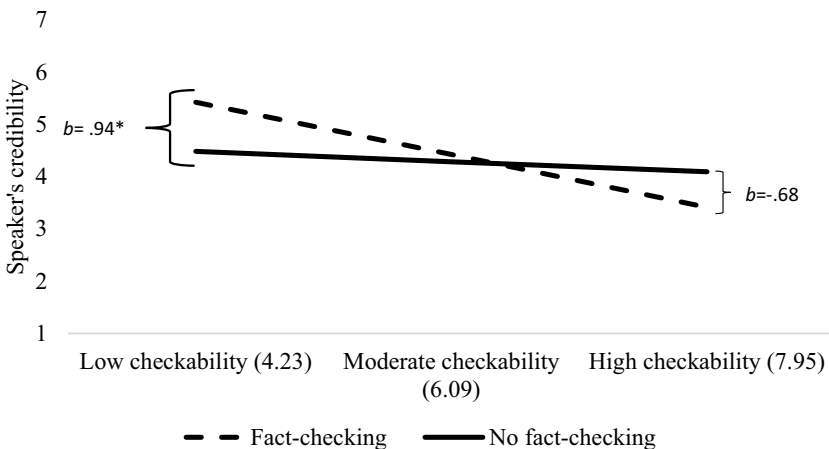

Figure 3. Estimated means of speaker's credibility at different levels of checkability detection (study 2).

Table 2. Direct and moderated effects of counter-attitudinal fact-checking (Study 2).

Predictors	Counter-attitudinal fact-checking				Pro-attitudinal fact-checking			
	Message accuracy		Speaker's credibility		Message accuracy		Speaker's credibility	
	Direct estimate (SE) 95% CI	Moderated estimate (SE) 95% CI	Direct estimate (SE) 95% CI	Moderated estimate (SE) 95% CI	Direct estimate (SE) 95% CI	Moderated estimate (SE) 95% CI	Direct estimate (SE) 95% CI	Moderated estimate (SE) 95% CI
Fact-checking	2.04* (.85) [.37, 3.72]		2.78* (.89) [1.03, 4.53]		-.80 (.94) [-2.66,1.05]		-.53 (.96) [-2.41,1.36]	
Checkab-ility	-.29** (.10) [-.48, -.09]	-.35* (.14) [-.61, -.08]	-.10 (.11) [-.31, .10]	-.43** (.14) [-.71, -.16]	-.19 (.10) [-.39,.01]	-.04 (.15) [-.33, .25]	-.10 (.10) [-.30,.10]	-.08 (.15) [-.38, .21]
$F_{(df1, df2)}$	19.98 (3, 183)***		11.88 (3, 183)**		9.26 (3, 176)***		7.65 (3, 176)**	
R^2	.25		.16		.14		.12	

Note. *$p < .05$, **$p < .01$, ***$p < .001$. The table presents unstandardized coefficients. The moderated estimates of checkability represent the interaction between fact-checking and message-checkability detection score.

Study 3: The motivations that underlie message-checkability detection

The goal of Study 3 was to test the specific mechanism that underlies message-checkability detection. Although Study 2 demonstrated that successful fact-checking might be contingent on the ability to discern fact- from opinion-based statements, there still remains considerable ambiguity regarding the antecedents of message-checkability detection. In particular, the literature points to, at least, two possible explanations: inaccuracy and motivated reasoning. To assess these explanations, Study 3 manipulated the type of motivation (accuracy vs. consistency) participants had while trying to classify fact- and opinion-based statements. Therefore, the final research question was:

RQ2: What is the effect of consistency (vs. accuracy) goals on the ability to distinguish fact- and opinion-based claims in value-laden messages?

Method

Design and participants
The current study tested the role played by processing goals (accuracy vs. consistency), with regard to message-checkability detection, and fact-checking. To this end, this study employed a between-subjects experimental design where participants were randomly assigned to three conditions that correspond with different processing motivations (accuracy, consistency, and control). Participants in the study ($N = 389$) were recruited from Qualtrics, applying equal quotas for political affiliation ($n_{Democrats} = 195$; $n_{Republicans} = 194$). The sample was predominantly female (242; 62.2%) and the average age was 56.04 ($SD = 14.89$).

Procedure and material
The procedure closely followed Study 2 with two notable differences. First, all participants were assigned to receive counter-attitudinal fact-checking (Republicans read a message that attempted to debunk misleading information in favor of gun rights and Democrats read a message that attempted to correct false information in favor of gun control). Second, all participants were instructed to complete a message-checkability detection task.

The manipulation of situational goals followed procedures utilized in previous studies (e.g., Kim, 2007). Specifically, prior to the message-checkability detection task and after exposure to the misinformation, participants in the consistency goals condition were reminded of their political affiliation ("Remember that you had selected …") and were told to be consistent with their political affiliation when making judgments about which

statements can be objectively verified ("With the goal of being as consistent as possible, would you consider each statement to be a factual statement or an opinion statement?"). Reminding participants of their political affiliation should prime their ideological identity and increase the degree to which they make decisions that are aligned with their preexisting beliefs (i.e., engage in motivated reasoning), prioritizing consistency over accuracy. In this case, participants primed with their political affiliation may be more likely to conclude that belief-consistent opinion statements are factual, and belief-inconsistent facts are non-verifiable opinions. Participants in the accuracy goal condition were not reminded of their political affiliation. Instead, they were told to be as accurate as possible when evaluating the statements ("With the goal of being as accurate as possible, would you consider each statement to be factual statement or an opinion statement?"). In line with previous studies, to encourage accuracy, *only* participants in the accuracy condition were told that the top three performers on the message-checkability detection task will receive a bonus of 40 USD. The control condition was exposed to the basic instructions of the task (the instruction utilized in Study 2) without mentioning consistency or accuracy.

Measures

The measures from Studies 1–2 were assessed, including *speaker's credibility* ($M = 2.81$, $SD = 1.75$, $\alpha = .95$), *message accuracy* ($M = 2.54$, $SD = 1.73$), and *message-checkability detection* ($M = 6.19$, $SD = 1.33$).

Results and discussion

Analysis of variance (ANOVA) with message-checkability detection score as the dependent variable and the situational goals manipulation as the independent variable tested the motivations that underlie the ability to discern fact- from opinion-based statements (RQ2). The results revealed a significant omnibus effect for situational goals on message-checkability detection; $F(2, 386) = 4.17$, $p = .016$, $\eta^2 = .02$. Tukey post-hoc test indicated that only the accuracy-goal condition significantly differed from the consistency-goal condition ($M_{diff} = .42$, $SE = .16$, $p = .029$) and the control condition ($M_{diff} = .40$, $SE = .16$, $p = .04$). Interestingly, there was no significant different between the consistency-goal condition and the control condition ($M_{diff} = -.02$, $SE = .16$, $p = .99$).

A mediation model (Model 4, 20,000 bootstrapped samples) in PROCESS (Hayes, 2018) tested the interplay between situational goals, message-checkability detection, and research outcomes. Situational goals were entered as a multicategorical predictor (with the control condition as a reference group), message-checkability detection as a mediator, as well as perceived accuracy and source's credibility as outcomes. The results revealed that processing information with accuracy-goals significantly increased message-checkability detection

($b = 0.40$, $SE = .16$, $p = .015$, 95% CI [.08, .72]), which in turn, reduced the perceived accuracy of the false information ($b = -.11$, $SE = .04$, $p = .048$, 95% CI [−.18, −.02]) and the credibility of its source ($b = -.13$, $SE = .07$, $p = .046$, 95% CI [−.26, −.01]). Notably, the models recorded a significant indirect effect of accuracy-goals on source credibility through checkability ($b = -.05$, $SE = .03$, 95% CI [−.14, −.01]). The direct effects on message accuracy, however, did not translate into an indirect effect through message-checkability detection ($b = -.04$, $SE = .03$, 95% CI [−.12, .01]).

General discussion

Research on "the effectiveness of fact-checking offers mixed results: some find that fact-checking reduces misperceptions, others that corrections are often ineffective" (Nieminen & Rapeli, 2019, p. 296). The primary objective of this paper directly addresses this discrepancy by proposing an underlying mechanism that explains why exposure to fact-checking appears to carry considerable effects under some circumstances while, in other cases, similar messages lack efficacy. Whereas the bulk of prior research had focused either on fact-checking message features (e.g., Amazeen et al., 2018; Garrett & Weeks, 2013; Young et al., 2018) or the characteristics of the audience (e.g., Jarman, 2016; Thorson, 2016; Weeks, 2015), the present inquiry attempted to provide a holistic account that examines the interplay between the two. In each experiment reported, pro-attitudinal fact-checking outperformed corrective information that ran counter to participants' ideological views. Ostensibly, the nonpartisan and independent nature of fact-checkers makes them ideal contenders to address political misinformation. Yet, as others have also observed (e.g., Jarman, 2016; Thorson, 2016; Walter et al., 2020), the robust effects of motivated reasoning, whereby individuals do not welcome information that contradicts their pre-established beliefs but show too great a readiness to believe things that align with their existing worldview, equally applies to fact-checking and to unverified information from partisan sources.

Extant research has documented the threats posed by motivated reasoning (Young et al., 2018). The present study extends this research by providing a more nuanced framework, not only explaining the discrepant findings surrounding fact-checking but also offering insight into how and why it occurs. In particular, in line with the implied truth effect (Pennycook et al., 2020), fact-checking of information that included both fact- and opinion-based statements proved to be especially challenging. More importantly, across two studies, the efficacy of fact-checkers to correct misperceptions was governed by participants' ability to accurately discern fact- from opinion-based claims. Put differently, exposure to counter-attitudinal fact-checking was able to root out previous misinformation only when participants were aware of which statements can be factually verified. This finding

is especially worrisome given that, on average, participants in the current (Studies 2–3) and in previous studies (Marpert et al., Pew Research Center, 2018a) show only limited ability to accurately determine whether political statements contain fact-based claims. Thus, the exclusive focus on fact-based statements that makes fact-checkers the "umpires of democracy" on paper (Lim, 2018, p. 1) can de facto impede their effectiveness. As argued by Uscinski and Butler (2013), part of the naïve epistemology of fact-checking is the tacit presupposition that "facts are unambiguous and not subject to interpretation" (p. 162). Yet, as our results suggest, even if fact-checkers can reach unambiguous verdicts regarding statements once they select them, the question what constitutes a verifiable claim will continue to be a major hurdle for healthy political discourse.

The original position of the current inquiry was that the limited ability to identify fact-based claims does not necessarily reflect motivated reasoning and it is plausible that, while trying to be accurate, people find it difficult to make such distinctions. Indeed, the task of identifying fact-based statements can be challenging as it involves being able to identify linguistic patterns and rules. Keeping in mind that individuals often use opinion-based claims to support their own positions (Kriplean et al., 2014), it stands to reason that people will be less sensitive to the use of opinion-based statements in the arguments of others. Inconsistent with this line of reasoning, however, the findings indicate that when primed with accuracy-goals, participants were able to significantly improve their ability to detect fact-based statements compared with those who were primed with consistency-goals and those in the control condition. More important, the performance of participants primed with consistency-goals did not differ significantly from their counterparts in the control condition. This suggests that, in the absence of motivation to be accurate, people default to misclassifying fact- and opinion-based claims. While it is unclear whether this is a result of a self-serving bias or a simple lack of motivation to engage with the message, it allows individuals to treat the misinformation and its correction more ambiguously, either by suggesting that not all fact-based statements were adequately addressed, or that some of the claims deemed false by the fact-checker are actually opinion-based statements. Ironically, it might be the case that while editors at fact-checking organizations are focused on "the art of finding a checkable fact" (Graves, 2013, p. 136), partisans are motivated to misclassify fact- and opinion-based statements.

The findings also carry practical implications for the future of fact-checking. The results suggest that there might be an additional step individuals take either before, during, or after they process a fact-checking message that focuses on value-laden beliefs. Namely, individuals identify the number of fact-based statements that can be checked and use it as a criterion for assessing the veracity of the entire message. After all, a statement that includes several imprecisions can still be relatively accurate if a large number of the fact-based claims are not

directly refuted. As a result, the fact-checking message and its verdict are not merely judged based on the information they address but also with regards to the presumed fact-based claims they fail to address. If individuals confuse opinions with facts, the task of fact-checking and debunking becomes even more complicated. One potential solution would be to change the unit of analysis for fact-checking and focus on individual statement as opposed to longer articles or speeches. This may eliminate or attenuate the need to identify fact-based statements and, as such, motivated reasoning may have less bearing on fact-checking. Alternatively, instead of reducing the unit of analysis, fact-checking messages can adopt a color code, or other ways of flagging fact- and opinion-based statements, such that individuals know what portion of the original article can be subjected to fact-checking (but see Eslick et al., 2011 for evidence that highlighting inaccuracies in a text can increase misinformation effects). In sum, the current investigation demonstrates that when ideologies clash, even the ability to ascertain fact-based claims becomes a thorny issue. Thus, irrespective of the specific solution, the future of effective fact-checking may be contingent on the willingness to distinguish fact- from opinion-based claims, regardless of who makes them and how they fit within our preexisting worldview.

Although this inquiry was designed to safeguard against various threats, it is not free of limitations. First, all three studies are limited to a single political issue. While the exclusive focus on gun policy allowed a close replication (Lindsay & Ehrenberg, 1993) in a highly polarized context, future research should analyze a variety of topics to examine the extent to which the same mechanism can be used to explain other categories of political misinformation. Moreover, following the procedure advocated in studies of motivated reasoning, the ideological-congruence variable was computed from different combinations of participants' political affiliations and the framing of the misinformation message (pro gun-control/pro gun-rights). Admittedly, the focus on political affiliation as a proxy of participants' support for gun policy may result in a less accurate measurement, yet it alleviates substantial threats concerning the pretest. Likewise, this decision is also justified given the intimate link between people's views on guns and their political affiliation (Silver, 2012).

Another caveat relates to the fact that we solely used fact-checking messages that debunked false information as opposed to fact-checking messages that confirmed true information. The efficacy of confirming (versus debunking) fact-checks in the presence of opinion-based statements may lead to different patterns of results. One possibility is that unaddressed opinions could be interpreted as false rather than unverified (Pennycook et al., 2020), which can negatively affect the efficacy of confirming fact-checking messages. Another possibility is that, given people's tendency to view unmarked information as true (e.g., Brashier & Marsh, 2020), the inclusion of opinion statements could *strengthen* the efficacy of confirming fact-checked messages, particularly when the claims are ideologically

congruent. Thus, the consequences of confirming fact-checking messages as well as fact-checks that both debunk and confirm should be interrogated in the future. Relatedly, the findings are also limited by our decision to focus on outright falsehoods ("Completely false") as opposed to more ambiguous scoring (e.g., "Half true" and "Mostly false"). The decision to focus on conclusive verdicts allowed for better internal control; rejection of the corrective information could be attributed to the participants as opposed to the ambiguity of the fact-check. At the same time, however, it also limited the external validity because much of political discourse, and subsequent fact-checking, is focused on the gray area between truths and lies.

Follow-up projects should also examine whether the ability to differentiate opinion- from fact-based claims improves when fact-checking messages focus on one specific claim at a time as opposed to several different statements in one fact-check. While the latter represents common scenarios where an hour-long speech or a political debate are being examined, fact-checking is typically concerned with a single statement. Arguably, when dealing with a single statement, the tendency to confuse fact- and opinion-based statements should carry less weight. Further, in the current project, we only consider short-term consequences, which limits the practical implications of our conclusions. It would be interesting to test whether the influence of fact-checking persists or, as with other corrective messages, people eventually revert back to inaccurate perceptions (Lewandowsky et al., 2012), particularly when the original message contains both opinion- and fact-based claims.

The current study also focused on the effects of exposure to fact-checking while neglecting an equally important question: What leads some people to seek out fact-checking information? There is some evidence to suggest that the public is very interested in fact-checking information (Stencel, 2016), which is somewhat counterintuitive considering the limited effect of such messages. It would be therefore informative to test whether people's ability to discern fact- from opinion-based statements and its subsequent effects on judgments is influenced by the extent to which individuals voluntarily choose to be exposed to this content. Lastly, as the ability to discern fact- from opinion-based statements appears to moderate the effectiveness of fact-checking messages, the next challenge would be to test if people can be trained to differentiate facts from opinions, and if so, whether this increases their sensitivity to fact-checking efforts.

Notes on contributors

Nathan Walter (Ph.D. University of Southern California, 2018) is an assistant professor at the Department of Communication Studies, Northwestern University. His research focuses on cognitive, metacognitive, and emotional processes at the heart of misinformation and its correction.

Nikita A. Salovich is a Ph.D. candidate in Cognitive Psychology in the Department of Psychology at Northwestern University. Her research examines when and why people are influenced by false information, and methods that support the evaluation of information.

Funding

References

Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, *31*(2), 211–236. https://doi.org/10.1257/jep.31.2.211

Amazeen, M. A. (2016). Checking the fact-checkers in 2008: Predicting political ad scrutiny and assessing consistency. *Journal of Political Marketing*, *15*(4), 433–464. https://doi.org/10.1080/15377857.2014.959691

Amazeen, M. A., Thorson, E., Muddiman, A., & Graves, L. (2018). Correcting political and consumer misperceptions: The effectiveness and effects of rating scale versus contextual correction formats. *Journalism & Mass Communication Quarterly*, *95*(1), 28–48. https://doi.org/10.1177/1077699016678186

Bennett, W. L., & Livingston, S. (2018). The disinformation order: Disruptive communication and the decline of democratic institutions. *European Journal of Communication*, *33*(2), 122–139. https://doi.org/10.1177/0267323118760317

Bode, L., & Vraga, E. K. (2015). In related news, that was wrong: The correction of misinformation through related stories functionality in social media. *Journal of Communication*, *65*(4), 619–638. https://doi.org/10.1111/jcom.12166

Brashier, N. M., & Marsh, E. J. (2020). Judging truth. *Annual Review of Psychology*, *71*(1), 499–515. https://doi.org/10.1146/annurev-psych-010419-050807

Chaiken, S., Liberman, A., & Eagly, A. H. (1989). Heuristic and systematic information processing within and beyond the persuasion context. In J. S. Uleman & J. A. Bargh (Eds.), *Unintended thought* (pp. 212–252). Guilford Press.

Clayton, K., Blair, S., Busam, J. A., Forstner, S., Glance, J., Green, G., Kawata, A., Kovvuri, A., Martin, J., Morgan, E., Sandhu, M., Sang, R., Scholz-Bright, R., Welch, A. T., Wolff, A. G., Zhou, A., & Byhan, B. (2020). Real solutions for fake news? Measuring the effectiveness of general warnings and fact-check tags in reducing belief in false stories on social media. *Political Behavior*, *42*(4), 1073–1095. https://doi.org/10.1007/s11109-019-09533-0

Druckman, J. N., & McGrath, M. C. (2019). The evidence for motivated reasoning in climate change preference formation. *Nature Climate Change*, *9*(2), 111–119. https://doi.org/10.1038/s41558-018-0360-1

Ecker, U. K. H., Lewandowsky, S., Cheung, C. S. C., & Maybery, M. T. (2015). He did it! she did it! No, she did not! Multiple causal explanations and the continued influence of misinformation. *Journal of Memory and Language*, *85*, 101–115. https://doi.org/10.1016/j.jml.2015.09.002

Egelhofer, J. L., & Lecheler, S. (2019). Fake news as a two-dimensional phenomenon: A framework and research agenda. *Annals of the International*

Communication Association, *43*(2), 97–116. https://doi.org/10.1080/23808985. 2019.1602782

Eslick, A. N., Fazio, L. K., & Marsh, E. J. (2011). Ironic effects of drawing attention to story errors. *Memory*, *19*(2), 184–191. https://doi.org/10.1080/09658211.2010. 543908

FactCheck.org. (n.d.). *Our mission*. https://www.factcheck.org/about/our-mission/

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). GPower 3: A flexible statistical power analysis program for the social, behavioural, and biomedical sciences. *Behavioural Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/ BF03193146

Flanagin, A., & Metzger, M. (2000). Perceptions of internet information credibility. *Journalism and Mass Communication Quarterly*, *77*(3), 515–540. https://doi.org/ 10.1177/107769900007700304

Garrett, R. K., & Weeks, B. E. (2013, February). *The promise and peril of real-time corrections to political misperceptions*. Proceedings of the 2013 Conference on computer supported cooperative work (pp. 1047–1058). Association for Computing Machinery, San Antonio, TX.

Grant, M. (1995). *Greek and Roman historians: Information and misinformation*. Routledge.

Graves, L. (2013). *Deciding what's true: Fact-checking journalism and the new ecology of news* [Unpublished dissertation], Graduate School of Journalism, Columbia University.

Graves, L. (2018). Boundaries not drawn: Mapping the institutional roots of the global fact-checking movement. *Journalism Studies*, *19*(5), 613–631. https://doi. org/10.1080/1461670X.2016.1196602

Graves, L., Gunther, A. C., Pelled, A., Su, M., Wang, Y., & Zhang, Y. (2018, May). *Effects of fact checks on partisan beliefs and perceptions of bias* [Paper presented]. At the annual meeting of the international communication association, Prague, Czech Republic.

Hayes, A. F. (2018). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach* (2nd ed.). Guilford Publications.

Jamieson, K. H. (2018). *Cyberwar: How Russian hackers and trolls helped elect a president. What we don't, can't, and do know*. Oxford University Press.

Jarman, J. W. (2016). Influence of political affiliation and criticism on the effectiveness of political fact-checking. *Communication Research Reports*, *33*(1), 9–15. https://doi.org/10.1080/08824096.2015.1117436

Kim, Y. M. (2007). Investigating the moderating effects of situational information processing goals in issue publics' Web behavior. *Communication Research*, *34*(2), 185–211. https://doi.org/10.1177/0093650206298069

Kriplean, T., Bonnar, C., Borning, A., Kinney, B., & Gill, B. (2014). *Integrating on-demand fact-checking with public dialogue*. Proceedings of the 17th ACM Conference on computer supported cooperative work & social computing - CSCW '14, 1188–1199, Baltimore, MD. https://doi.org/10.1145/2531602.2531677

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, *108*(3), 480–498. https://doi.org/10.1037/0033-2909.108.3.480

Kuru, O., Pasek, J., & Traugott, M. W. (2017). Motivated reasoning in the perceived credibility of public opinion polls. *Public Opinion Quarterly*, *81*(2), 422–446. https://doi.org/10.1093/poq/nfx018

Lewandowsky, S., Ecker, U. K., Seifert, C. M., Schwarz, N., & Cook, J. (2012). Misinformation and its correction continued influence and successful debiasing.

Psychological Science in the Public Interest, *13*(3), 106–131. https://doi.org/10.1177/1529100612451018

Lim, C. (2018). Checking how fact-checkers check. *Research & Politics*. Advance online publication, *5*(3), 1–7. https://doi.org/10.1177/2053168018786848

Lindsay, R. M., & Ehrenberg, A. S. C. (1993). The design of replicated studies. *American Statistician*, *47*(3), 217–228. https://doi.org/10.1080/00031305.1993.10475983

Margolin, D. B., Hannak, A., & Weber, I. (2018). Political fact-checking on Twitter: When do corrections have an effect? *Political Communication*, *35*(2), 196–219. https://doi.org/10.1080/10584609.2017.1334018

Maslow, A. (1954). *Motivation and personality*. Harper.

Merpert, A., Furman, M., Anauati, M. V., Zommer, L., & Taylor, I. (2018). Is that even checkable? An experimental study in identifying checkable statements in political discourse. *Communication Research Reports*, *35*(1), 48–57. https://doi.org/10.1080/08824096.2017.1366303

Nieminen, S., & Rapeli, L. (2019). Fighting misperceptions and doubting journalists' objectivity: A review of fact-checking literature. *Political Studies Review*, *17*(3), 296–309. https://doi.org/10.1177/1478929918786852

Pennycook, G., Bear, A., Collins, E. T., & Rand, D. G. (2020). The implied truth effect: Attaching warnings to a subset of fake news stories increases perceived accuracy of stories without warnings. *Management Science*, *66*(11), 4944–4957. https://doi.org/10.1287/mnsc.2019.3478

Pew Research Center. (2018a). *Distinguishing between factual and opinion statements in the news*. The Pew Charitable Trusts. http://www.journalism.org/2018/06/18/distinguishing-between-factual-and-opinion-statements-in-the-news/

Pew Research Center. (2018b). *7 facts about guns in the U.S. Retrieved July 2019*. The Pew Charitable Trusts. https://www.pewresearch.org/fact-tank/2018/12/27/facts-about-guns-in-united-states/

PolitiFact. (2019). *Statements about guns*. Poynter Institute. https://www.politifact.com/subjects/guns/

Schaffner, B. F., & Luks, S. (2018). Misinformation or expressive responding? *Public Opinion Quarterly*, *82*(1), 135–147. https://doi.org/10.1093/poq/nfx042

Silver, N. (2012 December). *Party identity in a gun cabinet*. ABC News Internet Ventures. https://fivethirtyeight.com/features/party-identity-in-a-gun-cabinet/

Stencel, M. (2016). *Public radio listeners want more fact-checking in election coverage*. Duke Reporters' LAB. https://reporterslab.org/public-radio-listeners-want-fact-checking-election-coverage/

Taber, C. S., & Lodge, M. (2006). Motivated skepticism in the evaluation of political beliefs. *American Journal of Political Science*, *50*(3), 755–769. https://doi.org/10.1111/j.1540-5907.2006.00214.x

Tenenbaum, J. B., Griffiths, T. L., & Kemp, C. (2006). Theory-based Bayesian models of inductive learning and reasoning. *Trends in Cognitive Sciences*, *10*(7), 309–318. https://doi.org/10.1016/j.tics.2006.05.009

Thorson, E. (2016). Belief echoes: The persistent effects of corrected misinformation. *Political Communication*, *33*(3), 460–480. https://doi.org/10.1080/10584609.2015.1102187

Uscinski, J. E., & Butler, R. W. (2013). The epistemology of fact checking. *Critical Review*, *25*(2), 162–180. https://doi.org/10.1080/08913811.2013.843872

Walter, N., Cohen, J., Holbert, R. L., & Morag, Y. (2020). Fact-checking: A meta-analysis of what works and for whom. *Political Communication*, *37*(3), 350–375. https://doi.org/10.1080/10584609.2019.1668894

Weeks, B. E. (2015). Emotions, partisanship, and misperceptions: How anger and anxiety moderate the effect of partisan bias on susceptibility to political misinformation. *Journal of Communication*, *65*(4), 699–719. https://doi.org/10.1111/jcom.12164

Weeks, B. E., & Gil de Zúñiga, H. (2021). What's next? Six observations for the future of political misinformation research. *American Behavioral Scientist*, *65*(2), 277–289. https://doi.org/10.1177/0002764219878236

Wood, T., & Porter, E. (2019). The elusive backfire effect: Mass attitudes' steadfast factual adherence. *Political Behavior*, *41*(1), 135–163. https://doi.org/10.1007/s11109-018-9443-y

Young, D. G., Jamieson, K. H., Poulsen, S., & Goldring, A. (2018). Fact-checking effectiveness as a function of format and tone: Evaluating FACTCHECK.org and FlackCheck.org. *Journalism & Mass Communication Quarterly*, *95*(1), 49–75. https://doi.org/10.1177/1077699017710453