ORIGINAL ARTICLE

Undermining the Corrective Effects of Media-Based Political Fact Checking? The Role of Contextual Cues and Naïve Theory

R. Kelly Garrett[1], Erik C. Nisbet[1], & Emily K. Lynch[2]

1 School of Communication, Ohio State University, Columbus, OH 43210, USA
2 Department of Political Science, Ohio State University, Columbus, OH 43210, USA

Media-based fact checking contributes to more accurate political knowledge, but its corrective effects are limited. We argue that biographical information included in a corrective message, which is often unrelated to the inaccurate claim itself, can activate misperception-congruent naïve theories, increasing confidence in a misperception's plausibility and inducing skepticism toward denials. Resistance to corrections occurs regardless of initial belief accuracy, but the effect is strongest among those who find the contextual information objectionable or threatening. We test these claims using an online survey-embedded experiment (N = 750) conducted in the wake of the controversy over the proposed Islamic cultural center in New York City near the site of the 9/11 attacks, and find support for our predictions. Theoretical and practical implications are discussed.

Political misperceptions are surprisingly resilient. Detailed reporting based on thorough research is not always enough to unseat inaccurate political ideas, as people are able to maintain false beliefs in the face of seemingly incontrovertible evidence (e.g., see Berinsky, 2012; World Public Opinion, 2006). In some circumstances, denials and debiasing strategies can even strengthen misperceptions' hold (Nyhan & Reifler, 2010; Sanna & Schwarz, 2003). Furthermore, a belief does not have to be deeply entrenched for this to occur: Even tentative, hastily formed, and weakly grounded beliefs tend to persevere (Anderson, Lepper, & Ross, 1980). Yet for democracy to succeed, a society must have the means to correct political misperceptions among its citizens. Failing this, there is significant risk that individuals will be left holding positions that they might otherwise abandon, or advocating for policies inconsistent with their own interests (Gilens, 2001; Lau, Andersen, & Redlawsk, 2008).

Misperceptions have many sources. Some are the result of individual cognitions, as when someone misremembers a fact, while others are grounded in communication.

Corresponding author: R. Kelly Garrett; e-mail: garrett.258@osu.edu

For example, people can be swayed by rumors—unverified claims made without strong evidence—and by misinformation—false information presented as truth and often accompanied by (seemingly) compelling evidence (DiFonzo & Bordia, 2007; Harrington, 2009). Providing factual information is an obvious response to political misperceptions (Jamieson & Gottfried, 2010; Kuklinski, Quirk, Jerit, Schwieder, & Rich, 2000), potentially inducing a more accurate state of political knowledge among those who encounter it. Although total accuracy and complete certainty may be the ideal, they are not always achievable in practice. Instead, fact-checking efforts that systematically evaluate the veracity and correctness of public statements, which were so prevalent in recent election cycles (e.g., FactCheck.org and Politifact), may be described as having corrective effects, whereby recipients' views become more closely aligned with empirical evidence or an individual becomes more confident in beliefs that are accurate (see, Lord, Lepper, & Preston, 1984).

This study extends research on debiasing and the continued influence of misinformation into the domain of political fact checking, examining how the content of a fact-checking message can influence recipients' willingness to accept the correction. Our contribution is twofold. First, we argue that the quality of the empirical evidence presented in the account is only one of several factors that individuals use when deciding what to believe. More specifically, we argue that contextual information unrelated to the inaccurate claim in any substantive way, such as photographs or biographical profiles commonly found in journalistic accounts, also influences recipients' response. This effect occurs when peripheral information brings to mind beliefs about the attitude object that are related to the targeted misperception, thereby increasing resistance to evidence included in the denial—even among those who initially held accurate beliefs. For example, an individual might be more likely to resist evidence that no weapons of mass destruction (WMDs) were found in Iraq if the fact-checking message were to mention the prevalence of anti-American sentiment in the Middle East than if it did not, even though the accuracy of the former does not depend on the latter. Invoking images of a region hostile to U.S. influence makes the idea that a country in that region would amass biological or nuclear weapons seem more plausible, even if it provides no concrete evidence. Second, we argue that individual characteristics shape these dynamics: The more objectionable the recipient finds the ideas activated by the contextual content, the stronger his or her resistance to the correction. Together, these claims suggest that contemporary journalistic conventions may inadvertently undermine corrective messages' effectiveness.

Naïve theories, stereotypes, and skepticism toward corrections
People regularly rely on a host of heuristics—mental shortcuts or guiding strategies that embody taken-for-granted principles and relationships—to make sense of the political world (Downs, 1957; Lupia & McCubbins, 1998; Popkin, 1991). Few individuals have the time or interest to carefully evaluate every piece of evidence relevant to a judgment task, be it a vote choice, candidate favorability, or issue position. Instead, humans constantly monitor the information environment for

patterns, spontaneously developing "naïve theories" about broad classes of objects to explain current observations and to make predictions about the future (Anderson & Lindsay, 1998). Naïve theories are a type of mental model that represent implicit beliefs and include causal elements, not just relational patterns. In contrast to schemata, these knowledge structures are often based on analogical thinking and tend to be more flexible, serving to represent unfamiliar, rather than routine, situations (Jones, Ross, Lynam, Perez, & Leitch, 2011). Instead of simply providing evidence that certain features co-occur, naïve theories simplify social perception by offering explanations for the co-occurrence. Individuals generally form these explanations based on top-of-the-head considerations and tend not to test them rigorously, yet they are powerful tools for making sense of a complex world. They are so useful, in fact, that they often outlive the evidence on which they were based (c.f., Seifert, 2002). For instance, when both naïve and scientific theories about the same concept are encountered, the scientific theory tends to suppress, but not replace, the naïve theory (Shtulman & Valcarcel, 2012).

A classic example of this phenomenon involved explanations brought to mind by *fabricated* evidence that either risk-taking (in one condition) or risk-averse (in the other) people make better firefighters (Anderson et al., 1980). Strikingly, participants tended to remain committed to a belief consistent with the "evidence" to which they were initially exposed even after the deception was revealed. This is not due to the fact that participants clung to belief in the initial report: The vast majority accepted that researchers had fabricated the evidence. Instead, it is attributed to the spontaneous creation of naïve theories. When these self-generated explanations were more salient, as when researchers explicitly asked participants to describe why risk-takers (or avoiders) make better firefighters, the beliefs were more robust.

This process presumably extends to claims about (false) political information. Even if evidence for a belief is discredited, the belief may be reinforced if compelling self-generated explanations for it are brought to mind by the correction (Seifert, 2002). The subsequent accessibility of these causal stories may further reinforce the belief because availability and ease of processing are frequently used heuristics for judging prevalence, certainty, and accuracy (Schwarz, Sanna, Skurnik, & Yoon, 2007; Tversky & Kahneman, 1973). Thus, any message that triggers thought about naïve theory will tend to produce greater confidence in theory-consistent beliefs, potentially undermining attempts at correction.

Humans' propensity to categorize others further contributes to this effect. When a trait category is activated, the observer is more likely to classify the person being observed in terms of the activated category, introducing bias and distortion into comprehension and storage processes (Higgins, Rholes, & Jones, 1977). For example, if an individual categorizes someone he or she encounters as "lazy," the observer is likely to attribute a variety of qualities stereotypically associated with laziness to the observed, such as slovenliness or apathy, even if these attributes are not immediately evident. We suggest that the observer in this scenario would also be more resistant to evidence that the individual observed did not exhibit these stereotypical attributes.

Individuals can retain and be influenced by *rejected* naïve theories as well, although these theories are unlikely to be as influential as those that are accepted (Anderson et al., 1980). For example, priming a racial category evokes many of the same cultural stereotypes—including naïve theories about race and behavior—regardless of the levels of prejudice the individual exhibits (see Macrae & Bodenhausen, 2000). The difference is that some individuals replace these stereotypic thoughts with nonprejudicial personal beliefs. In the context of fact checking, this suggests that a message may prompt consideration of stereotypical relationships that influence beliefs even if the stereotypes are themselves rejected. Returning to the WMD example, we would anticipate that priming the Muslim religious category in a fact-checking message makes it more likely the recipient would acknowledge the possibility of weapons in Iraq even if the individual rejects the stereotypical view that Islam is inherently violent.

Finally, when a misperception reflects negatively on the target, as is often the case in the political domain, naïve theories about the *threat* posed by the target greatly undercut any challenge to this belief. Naïve theories which imply that a group, whether it is defined by religion, race, ideology, or something else, threatens the dominant social order can be particularly powerful. Thus, contextual information such as a photograph or a biographical sketch that identifies a political actor as belonging to a threatening group activates relevant negative cognitions, makes related ideas more accessible, and makes the claim that the target is itself threatening more plausible. Ultimately, this is likely to increase skepticism toward any counterevidence. This prediction is consistent with evidence that people are more willing to accept claims that President Barack Obama is a socialist when race is salient (Kosloff, Greenberg, Schmader, Dechesne, & Weise, 2010). Our argument goes further, however, suggesting that priming thoughts about race would also make this view more difficult to unseat, especially among those who find racial minorities to be threatening.

Fact checking the "Ground Zero Mosque"

The controversy over the proposal to build an Islamic cultural center and mosque near the site of the 9/11 attacks in the fall of 2010 provides a unique opportunity to test these theoretical mechanisms in a real-world setting that evokes numerous challenges to American hegemonic values. Specifically, we tested a correction to the widespread belief that Feisal Abdul Rauf, the Imam most publicly associated with the project at the height of the controversy, refused to condemn terrorism. Many Americans harbor concerns about Islam, and most have encountered claims that it is more likely than other religions to embrace violence, that most Islamic religious leaders are fundamentalists who are prone toward extremism, and that Muslims living in America are not supportive of the United States (Gallup Poll, 2011; Nacos & Torres-Reyna, 2006; Nisbet, Ostman, & Shanahan, 2009; Pintak, 2006; Shaheen, 2001). Activating naïve theories about the relationship between the Islamic faith and religious imperialism, even if the theories only hold in certain circumstances or

are not fully accepted, would make claims about a specific Imam endorsing Islamic terrorism more plausible, increasing individuals' resistance to denials. And priming people to think about the "radical Muslim" category will lead people to think of the Imam as such, which can further increase denial resistance.

There are several kinds of information that, if included in a rebuttal, are expected to promote these types of cognitions. Consider, for example, that when condemning the 9/11 attacks, Mr. Abdul Rauf expressed his belief that the terrorist action, though unjustifiable, was motivated in part by harmful U.S. policies in the Middle East. The Imam was not alone in this view—many political commentators argued the 9/11 attacks were "blowback" from U.S. policies in the Middle East and Afghanistan (see Coll, 2004; Gibbs, 2005; Moore, 2004)—but the position clearly challenges hegemonic cultural conventions about how Americans view the terrorism. Some may also consider the motivations for making such a statement to be ambiguous. Spoken by a foreign policy expert, it would likely be seen as a legitimate attempt to understand the causes of violence; however, when spoken by an Islamic religious leader the same statement might be viewed as evidence of Islamic extremism. Consider also Mr. Abdul Rauf's advocacy for the integration of Muslims into mainstream U.S. society. This is a quintessential example of the America-as-melting-pot metaphor, but viewed in the context of fears about Islamic Fundamentalism, it could be seen as an ominous vision of Islamic world domination often promoted in American political and religious discourse (Nacos & Torres-Reyna, 2006; Nisbet et al., 2009; Pintak, 2006; Seib, 2005). In these ways, even information that is unrelated to the initial misperception could lead people to resist counterevidence.

Pictures are another common journalistic element that could have harmful consequences on political fact checking's corrective effects. Image processing is almost always fast and unreflective. Images encountered for even a fraction of a second can have a strong influence on impressions, contributing to a "gut feeling" about attributes such as competence, trustworthiness, or likability (Bar, Neta, & Linz, 2006; Hall, Goren, Chaiken, & Todorov, 2009). Unless individuals are strongly motivated to process the message in its entirety, these immediate reactions can function as peripheral cues (Petty & Cacioppo, 1986). Rather than weighing the cognitively complex textual evidence, many individuals will tend to use their gut reaction to images as a heuristic for how to respond to the message as a whole. Given the high levels of automaticity in these processes, images could also unconsciously promote other types of cognitions. For example, an image that conveys the Imam's Middle Eastern cultural heritage could activate theories about Islamic religious leaders as dangerous outsiders who hold radical beliefs in ways that an image of the Imam in Western-style attire would not. Factual details about the Imam's efforts to fight Islamic extremism would be inconsistent with such cognitions, and would therefore be less persuasive in the context of the former image than the latter. Even individuals who steadfastly reject negative stereotypes of Muslims would be more cognizant of the existence of extremists, and would therefore be more likely to acknowledge that the Imam *might* be a member of a radical minority when faced with the stereotype-evoking image.

The theoretical arguments presented thus far generate two testable predictions. We predict that *a fact-checking message will be less effective when it includes statements made by the Imam that challenge dominant U.S. cultural hegemony (H1a)*. We also assert that *a fact-checking message will be less effective when paired with an image highlighting the Imam's unique cultural heritage than with an image that is more compatible with dominant U.S. cultural hegemony (H1b)*.

Although we predict that contextual content will reduce the corrective effects for everyone, we also expect that information that activates naïve theories about Islamic extremism and about challenges to American social hierarchy will vary according to the attitudes and beliefs of the message recipient. When messages are processed peripherally message position is itself a cue, meaning that attitude-consistent messages will be more readily accepted than attitude-discrepant messages (Petty & Wegener, 1998). These biases will moderate the influence of the activated naïve theories, which will be expressed in at least two different ways. First, *the more objectionable someone finds beliefs that the US is itself in part culpable for the 9/11 attacks or that integration of Muslims into American society is a social good, the more likely Mr. Abdul Rauf's statements are to activate negative cognitions about Islamic extremism, and the less persuasive they will find the rebuttal to be (H2)*.

Second, the more invested the recipient of the rebuttal is in preserving or defending the existing social hierarchy, the more likely the individual is to object to either of these transgressions against American hegemony. The stable propensity to support existing group hierarchies and to prefer that one's own in-group maintain social dominance has been termed social dominance orientation (SDO) (Pratto, Sidanius, Stallworth, & Malle, 1994; Sidanius & Pratto, 1999). Individuals who are high in SDO prefer inequality among social groups and are more likely to hold attitudes and engage in behaviors that maintain existing social and cultural hegemonies, to actively oppose any changes that favor minority groups, and to take actions to preserve their own ingroup values when presented with symbolic or cultural threats (Morrison & Ybarra, 2009; Pratto et al., 1994; Sidanius & Pratto, 1999). For example, Thomsen et al. (2008) found that individuals who scored high on SDO were more willing to persecute members of minority groups (Muslims and Latinos) that were presented as assimilating into American culture than minority members that were portrayed as remaining segregated. Claims that U.S. policies toward the Middle East have been flawed, or that Muslim Americans should be an integral part of U.S. society, imply that social change is necessary, and they will be viewed more negatively as SDO increases.

As a consequence, *the higher an individual's social dominance orientation, the less effective a fact-checking message including statements challenging dominant U.S. cultural hegemony will be (H3a)*. Individuals who oppose changes to the existing social order will also respond negatively to more passive challenges of Western conventions, and therefore *the higher an individual's social dominance orientation, the less effective a fact-checking message paired with an image highlighting the Imam's unique cultural heritage will be (H3b)*.

Boomerang effect

A message designed to change attitudes, beliefs, or behaviors has the potential for a boomerang effect, whereby it induces a shift away from the intended outcome (Byrne & Hart, 2009). In politics, this can mean that negative information about a preferred candidate can lead people to view that candidate more favorably (Meffert, Chung, Joiner, Waks, & Garst, 2006; Redlawsk, 2002). Recent research suggests that factual corrections to political misperceptions are no exception. Faced with evidence that their beliefs are inaccurate, some individuals will embrace these beliefs more strongly (Hart & Nisbet, 2012; Nyhan & Reifler, 2010).

The mechanisms explaining this phenomenon can be classified in two broad groups (Byrne & Hart, 2009). *Intended construct activation* occurs when message recipients accurately comprehend a message, but are unable or unwilling to act on the information provided. For instance, fact-checking messages undermine individuals' ability to endorse claims that have been identified as false, which some may view as a threat to their intellectual autonomy. In a form of psychological reactance (Brehm, 1966; Byrne & Hart, 2009) these individuals may attempt to reassert their independence by embracing the (inaccurate) belief more strongly. This study, however, concerns *unintended construct activation*, which occurs when recipients focus on something other than the intended message. The common tendency to counterargue attitude-discrepant political messages (Lord, Ross, & Lepper, 1979; Taber & Lodge, 2006) means that exposure to fact checking can serve to activate attitude-consistent evidence, potentially reinforcing support for the disputed claim (Byrne & Hart, 2009; Nyhan & Reifler, 2010).

Focusing specifically on unintended, rather than intended, construct activation, this study provides another opportunity to examine under what conditions a boomerang effect may occur. The question here is *whether the text and images contained within the corrective message promote support for misperceptions (RQ1).*

Method

A demographically diverse sample ($N = 750$) of U.S. residents whose characteristics resemble the country's population was recruited from an opt-in online panel administered by Survey Sampling International. The sample is 45% male and includes a range of ages (44% 18−35 years, 26% 36−50 years, 30% older), education levels (23% high school or less, 40% some college, 37% four years college or more), and incomes (60% <$50K, 21% $75K+). Importantly, the sample also includes a range of religious affiliations (44% Protestant, 20% Catholic, 22% Atheist/Agnostic, 14% other), political affiliations (40% Democrat, 28% Independent, 32% Republican), and ideologies (30% liberal, 27% moderate, 33% conservative). Only two Muslims were included in the sample, which is not surprising as they represent less than 1% of the American population and are geographically concentrated (Pew Forum on Religion & Public Life, 2012). These individuals were excluded from all analyses as the

study concerns potential biases in perceptions of Muslim Americans. The diversity of our sample enhances the external validity of this experimental study.

Procedure

Data were collected via an experiment that was embedded within an online public-opinion survey and that utilized a pretest–posttest control group design. The survey was conducted between September 14 and September 19, 2010, in the wake of August protests over the proposal to build an Islamic cultural center and mosque approximately two blocks from the site of the 9/11 attacks, just days after the *New York Times* published an op-ed about the project authored by Mr. Abdul Rauf (September 7, 2010). Participants began by answering questions about their media habits, SDO, attitudes toward Muslims and the Islamic cultural center, including their agreement with a pair of controversial ideas about Muslims and the United States, and their familiarity with four widely discussed misperceptions related to the controversial project. Of particular concern to this study was belief in the statement that "Feisal Abdul Rauf, the Imam backing the proposed Islamic cultural center and mosque, is a terrorist-sympathizer who refuses to condemn Islamic attacks on civilians." There was substantial evidence at the time of the study that this statement was false according to FactCheck.org, a nonpartisan service run by the Annenberg Public Policy Center, and Politifact, the Pulitzer-prize winning service of the *St. Petersburg Times*.

Participants first indicated the frequency with which they encountered the statement and its counterclaims. Almost half (45%) had prior contact with the false claim, and although most reported hearing it only infrequently, about one in nine participants (11%) indicated they had heard it "many times." Only about one in six (16%) had encountered a rebuttal of this claim, which is just over a third (36%) of those who had previously encountered it. Participants then assessed the statement's accuracy on a 5-point scale: definitely true, probably true, probably false, definitely false, or by indicating that they were "unsure." Going into the study, slightly more participants rejected the claim than believed it (38% vs. 32%), and the remainder were unsure. This variable was recorded so that higher values correspond to greater belief in the statement (*Range* $= 0–4$, $M = 1.89$, $SD = 1.15$).

The experimental manipulation that followed varied the form of the rebuttal to the false statement about the Imam. We programmed the survey software to randomly assign participants to one of five conditions, with balancing to ensure comparable numbers were assigned to each condition. The rebuttal in the first condition consisted of a three-paragraph, 198-word text-only message telling participants that FactCheck had reviewed the evidence and concluded that the statement was false (see Appendix for complete wording). The message also included several pieces of information that directly contradicted the inaccurate claim. The second condition was based on the first, but added two factually accurate, but potentially controversial, statements suggesting that the Imam was a threat to U.S. social hierarchy and that emphasized his Muslim American identity. These statements were appended to the end of the

second and third paragraphs of the message, respectively: "Although Mr. Adbul Rauf has said that the U.S. bears responsibility for harm caused by its policies toward the Middle East, he is quick to point out that terrorism is never justified" and "Our mission is to interweave America's Muslim population into mainstream society. We are a Muslim-American force for promoting the universal values of justice and peaceful coexistence . . . "

The third and fourth conditions used the low-controversy text from the first condition, but added one of two color photographs, presented in the upper-left corner of the frame containing the fact-checking message. One photograph, taken in New York City, showed the Imam dressed in a Western-style business suit, standing behind a podium, surrounded by a racially diverse crowd of similarly dressed adults (see Appendix for both images). The other image was taken at a mosque in Bahrain, as the Imam greeted worshipers after leading midday prayers. The four most prominent men in the picture are dressed in white, and three of them wear keffiyehs, the traditional Arab headdress. Although the Imam's head is not obviously covered, he also wears a bisht, a traditional black cloak with gold trim common in the Middle East.

In the fifth condition, the control, no rebuttal was presented. Participants were routed directly to the follow-up questionnaire.

After the manipulation, participants were asked to assess the false statement a second time "based on everything you know" ($M = 1.67$, $SD = 1.22$). Overall, almost half (49%) the respondents rejected the statement at this point, though the number who believed it was almost unchanged (30%). In other words, only the number of people responding that they were "unsure" declined. The survey questionnaire concluded by asking participants a battery of sociodemographic questions, including questions about ideology and religious beliefs.

Measures

We utilized a well-established four-item summative index to measure SDO (Pratto et al., 1994), which included one reverse-coded item ($M = 10.47$, $SD = 4.04$, $\alpha = .66$). Although the reliability of this measure is at the margin of acceptability, it has been extensively validated and is widely used in the field. Attitudes toward the controversial ideas included in some versions of the fact-checking message were measured prior to exposing participants to the fact-checking messages, using a pair of 6-point scale items anchored by strongly disagree (1) and strongly agree (6). The statements were "Muslims enrich and contribute to American society" ($M = 3.74$, $SD = 1.42$) and "The 9/11 attacks were in part a response to harmful U.S. policies in the Middle East" ($M = 3.33$, $SD = 1.50$).

Ideology was measured on a seven-point scale anchored by "very liberal" and "very conservative," with higher values corresponding to greater conservatism ($M = 4.04$, $SD = 1.64$). Education was also measured on a 7-point scale, and was reverse coded so that higher numbers denote greater educational attainment. The modal value of this variable was five, corresponding to "some college" ($M = 4.94$, $SD = 1.37$). Fox News use was included in a battery of 16 media-exposure measures; participants indicated

how much they relied on each outlet, on television or online, using a 7-point scale anchored by "never" and "a great deal" ($M = 3.05$, $SD = 2.21$). A variety of other outlets were tested in the model, but only Fox News had an effect, and so the other outlets are excluded from these analyses. Finally, three multiple-choice items tapped knowledge about Islam by asking participants to name the largest sect within Islam, the day of the week that most Muslims gather to pray, and the city to which Muslims face when they pray. The variable was constructed by counting the number of correct answers given ($M = 1.03$, $SD = .80$).

This last section of the questionnaire also included a five-item social desirability measure, which we used to assess whether this characteristic produced systematic bias in beliefs about the Imam. Participants were asked, on a three-point scale (no, unsure, or yes), whether they had ever hurt or taken advantage of another person, were always good listeners, were always courteous, and were always willing to admit their mistakes. The more strongly a participant denied these unavoidable human weaknesses, the higher the social desirability score ($M = 11.53$, $SD = 2.60$). The factor was not a significant predictor in any of the models reported here and is therefore excluded from further consideration.

Results

We begin with a brief overview of the effects of the message manipulation. Table 1 presents a summary of the shifts in participant attitudes across the various conditions. Almost three-quarters of the participants in the control condition saw no change in their beliefs, and individuals who did express a different view were fairly evenly divided between those whose accuracy increased and those whose accuracy decreased. Results in the other conditions were generally encouraging, suggesting a disproportionate shift in attitude toward greater accuracy after exposure to the correction.

More rigorous analyses confirms that the fact-checking messages presented in this study reduced the extent to which participants thought the false statement about the Imam was true in three out of four conditions. Only the rebuttal containing the controversial text was ineffective when compared to the no-rebuttal condition. Evidence for this conclusion comes from a linear regression model predicting

Table 1 Proportion of Participants Changing Belief by Condition

Condition	Less accurate (%)	No change (%)	More accurate (%)
No rebuttal (control)	15.8	70.4	13.8
Text only: no controversial statements	11.9	53.0	35.1
Text only: controversial statements	15.7	51.6	32.7
Photo: Imam in Western-style attire	8.8	54.7	36.5
Photo: Imam in Middle-Eastern-style attire	16.3	49.0	34.7

Table 2 Predicting Acceptance of False Statement by Rebuttal Condition

	Model 1	Model 2	Model 3
Rebuttal conditions[a]			
Text only: no controversial statements	-0.41^{***} (0.10)	-0.44^{***} (0.10)	-0.43^{***} (0.10)
Text only: controversial statements	-0.18 (0.10)	-0.18 (0.10)	-0.16 (0.10)
Photo: Imam in Western-style attire	-0.54^{***} (0.10)	-0.56^{***} (0.10)	-0.55^{***} (0.10)
Photo: Imam in Middle-Eastern-style attire	-0.29^{**} (0.10)	-0.29^{**} (0.10)	-0.30^{**} (0.10)
Objection to ideas in controversial statements	0.06^{***} (0.02)	0.09^{**} (0.03)	0.10^{**} (0.03)
No controversial statements × Objection	—	-0.11^{*} (0.05)	-0.12^{*} (0.05)
Controversial statements × Objection	—	-0.03 (0.05)	-0.06 (0.05)
Western-style attire × Objection	—	0.04 (0.04)	0.03 (0.04)
Middle-Eastern-style attire × Objection	—	-0.06 (0.05)	-0.07 (0.05)
SDO	0.03^{***} (0.01)	0.03^{***} (0.01)	-0.00 (0.02)
No controversial statements × SDO	—	—	0.04 (0.03)
Controversial statements × SDO	—	—	0.07^{*} (0.03)
Western-style attire × SDO	—	—	0.01 (0.03)
Middle-Eastern-style attire × SDO	—	—	0.05^{*} (0.03)
Controls:			
Belief in false statement prior to rebuttal	0.43^{***} (0.04)	0.43^{***} (0.03)	0.43^{***} (0.03)
Exposure to statement prior to experiment	0.15^{***} (0.04)	0.15^{***} (0.04)	0.16^{***} (0.04)
Exposure to rebuttals prior to experiment	-0.18 (0.10)	-0.21^{*} (0.10)	-0.21^{*} (0.10)
Ideology (more conservative)	0.05^{*} (0.02)	0.05^{*} (0.02)	0.05^{*} (0.02)
Education	-0.04 (0.03)	-0.04 (0.03)	-0.04 (0.03)
Knowledge about Islam	-0.12^{**} (0.04)	-0.12^{**} (0.04)	-0.12^{**} (0.04)
Extent of Fox News use	0.08^{***} (0.02)	0.08^{***} (0.02)	0.08^{***} (0.02)
Observations	743	743	743
R^2	0.487	0.494	0.500

Standard errors in parentheses. Statement acceptance, the predicted value, was measured on a 5-point scale where higher values correspond to greater acceptance.

[a]Control condition (no rebuttal) is reference category.

$^{*}p < .05,\ ^{**}p < .01,\ ^{***}p < .001.$

misperception acceptance as a product of condition, after controlling for several other influential factors (see Table 2, model 1). As would be expected, initial belief in the misperception was the strongest predictor of postexposure belief. (This is the preferred method for modeling a change in belief, rather than employing a single change-score variable, as it avoids the assumption that the difference between pre- and postmeasures is fixed and removes the problem of overcorrecting the postscore in the analysis (Cohen, Cohen, West, & Aiken, 2003, see pp. 570–571 for discussion).)

The frequency of exposure to the false statement prior to beginning the experiment was also a positive predictor, though its effect was smaller. The overall explanatory power of this model is high, capturing almost half the variance. Although this is due in large part to the control variables, estimating the model in stages confirms that adding the predictors of interest as a block after entering the controls produces a significant improvement in model fit, *change in* $R^2 = 0.043$, $F(6,729) = 10.11$, $p < .001$ (not shown in table).

The coefficients on all four conditions are negative and three were significant, with *b*s between -0.28 and -0.54, all *p*s $< .01$. The effect of the more controversial

text-only rebuttal was smaller than the rest, rendering it nonsignificant, $b = -0.18$, $p = .9$. To more fully appreciate the substantive significance of these results we also estimated an analysis of covariance (ANCOVA) and plotted the estimated marginal means of belief in the false statement (and corresponding 95% confidence intervals) for each of the five conditions when holding all other variables at their mean (see Figure 1). Viewed in this way, we see that a typical participant exposed to the fact-checking message was inclined to reject the false claim, whereas participants who were not exposed to it tended to be unsure of its truth status.

Turning to our hypotheses, the data support our first prediction, namely that a rebuttal including statements by the Imam that challenge the dominant cultural hegemony would be less effective than one which did not include those statements (H1a). The coefficient on the condition that included the controversial information in the rebuttal is of significantly smaller magnitude than that of the less controversial text-only condition, $\chi^2(1, n = 736) = 5.84$, $df = 1$, $p < .05$ (Bonferroni adjusted p-value). Graphically, we see in Figure 1 that the bar corresponding to the text-only message without the controversial statements is shorter than the bar for the text-only message with controversial content.

Our second prediction was that a rebuttal paired with an image of the Imam that highlights his Middle-Eastern cultural heritage would be less effective than one
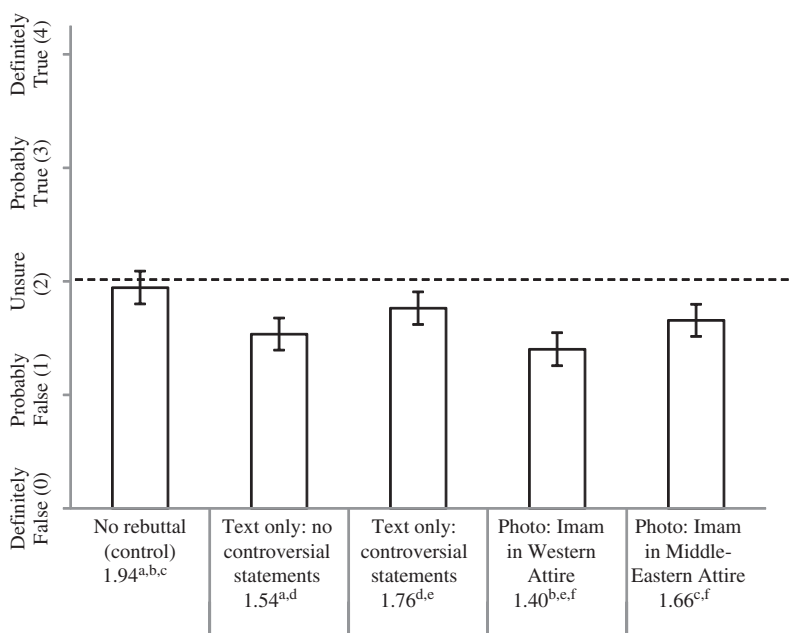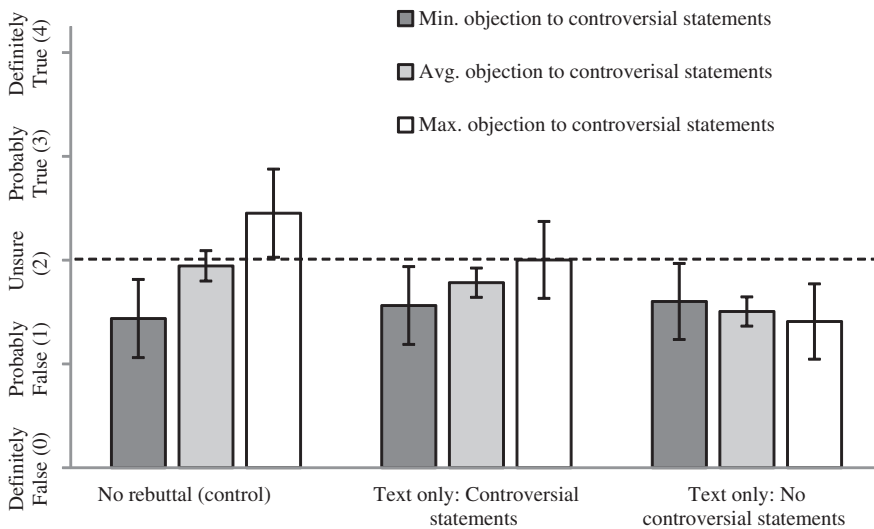

Figure 1 Predicted acceptance of false statement by rebuttal condition.
Note: Estimated marginal means holding other variables at their mean. Lower values are more accurate. Error bars represent 95% confidence intervals. Matching superscripts denote differences significant at $p < .05$ based on ANCOVA results.

with an image in which he embodies a more mainstream vision of American identity (H1b). This prediction was also supported by the data: The image showing the Imam in Western-style clothing was associated with a significantly larger drop in belief than the image showing the Imam dressed in a Middle-Eastern style, $\chi^2(1, n = 736) = 5.41$, $p < .05$ (Bonferroni adjusted p-value). Graphically, the bar in Figure 1 associated with Western-style attire is shorter than the bar associated with Middle-Eastern-style attire.

Variations in effects associated with individual attributes
Next we consider how individual-level differences shape participants' response to the different versions of the rebuttal. We hypothesized that the rebuttal that included the Imam's controversial statements would be less effective the more the recipient objected to these sentiments (H2). A regression model examining this moderating relationship, however, suggests a slightly different pattern (see Table 2, model 2). According to this model, strength of objection to the Imam's statements is positively correlated with belief in the false statement about him in every condition, $b = 0.10$, $p < .01$ except the low-controversy, text-only rebuttal. In that one case, the influence of objection to the statements is not significantly different than zero. Figure 2 helps to illustrate this relationship. In the no-rebuttal condition, we can see that a participant who objected strenuously to the ideas that U.S. policy may have helped to motivate the 9/11 attacks or that Muslims contribute to U.S. society was more likely to believe that the Imam was a terrorist sympathizer than one who did not object to these ideas. The pattern observed when participants were presented with the text-only

Figure 2 Predicted acceptance of false statement by rebuttal condition and participant objection to controversial statements used in rebuttal.
Note: Estimated marginal means holding other variables at their mean. Lower values are more accurate. Error bars represent 95% confidence intervals.

fact-checking message that included the Imam's statements about these controversial topics was not significantly different. However, when the fact-checking text did not include the controversial statements, objection to these counterhegemonic ideas no longer promoted misperception.

The final two hypotheses concern the moderating role of SDO (H3a and H3b). The theoretical model presented here led us to expect that individuals high on SDO would find the rebuttals containing counterhegemonic information to be less persuasive, and the results are consistent with these expectations. The third block of interaction terms (Table 2, model 3) indicates that SDO significantly moderates the influence of the rebuttal containing the controversial statements made by the Imam, $b = 0.07$, $p < .05$, and the rebuttal paired with an image of the Imam dressed in a Middle-Eastern style, $b = 0.05$, $p < .05$. In both cases, the corrective effects of the rebuttals were lower among those with higher SDO.

No evidence of the boomerang effect

Finally, we also posed one research question in our study. We wondered whether inclusion of the counterhegemonic contextual information would activate unintended constructs that would produce a boomerang effect whereby participants' beliefs moved away from the factual information presented in the rebuttal and toward the misperception. The answer based on these data is no. Returning to examine Table 2 (and Figure 1) we see that none of the four conditions systemically promoted inaccurate beliefs. To the contrary, only one condition failed to significantly increase accuracy despite a large sample and an observed power of above 0.99 (computed post hoc based on the variance explained by adding condition).

This test is, however, incomplete; prior studies have suggested that backfire effects are limited to individuals who are most inclined to resist the correction (Nyhan & Reifler, 2010). Thus, perhaps the boomerang effect will only be evident among the strongest partisans, ideologues, or critics of the Imam. To examine this possibility, we constructed several regression models, each testing a different potential moderator. There was no evidence of a significant interaction between the treatment conditions and party, ideology, or attitudes and emotions related to the proposed Islamic cultural center, such as anger, fear, or the project's perceived costs. Thus, we have no evidence that the background information included in a fact-checking document increased belief in the claims being challenged.

Discussion

The results of this experiment demonstrate that the effectiveness of a correction to an inaccurate political belief depends not just on the strength of the relevant evidence, but also on other, less relevant information included in the message. In this study, both textual and photographic information with no direct bearing on the false statement reduced a message's corrective effects. The unfortunate conclusion that we draw from this work is that contextual information so often included in fact-checking messages

by professional news outlets in order to provide depth and avoid bias can undermine a message's corrective effects. We suggest that this occurs when the factually accurate information (which has only peripheral bearing on the misperception) brings to mind naïve theories that are consistent with the misperception. Once activated, processing the misperception in these terms becomes easier, increasing confidence in its plausibility and inducing skepticism toward denials. Both text and images can produce these effects, and the activated heuristic does not need to be strongly believed or to have direct bearing on the misperception itself.

The influence of the contextual cues on the response to a rebuttal is also partly a function of the recipient's disposition. In this study, we focused on a misperception concerning a Muslim-American religious leader who backed a controversial proposal to build an Islamic cultural center near the site of the 9/11 terrorist attacks. Although Mr. Abdul Rauf has spoken openly against religious extremism, he clearly identifies himself with his faith, has repeatedly expressed his belief that Muslim Americans should be an integral part of U.S. society, and has criticized the US for its policies in the Middle East. By these actions, he is implicitly challenging the dominant social order of the US that emerged in the wake of the 2001 terrorist attacks. As a consequence, the more oriented an individual is toward defending the existing hierarchy, the more resistant he or she will be to rebuttals that inadvertently activate naïve theories about the threat posed by Islam and by Islamic religious leaders.

We failed to find evidence of a boomerang effect, regardless of the presence or absence of counterhegemonic contextual information. The absence of an effect associated with background information is perhaps unsurprising: Compared to the effects of the fact-checking message itself, the effects of this more peripheral information are expected to be small. The bigger question is why we found *no* systematic evidence of individuals embracing a misperception as a result of their exposure to the rebuttal despite our selection of a contentious and highly polarized political issue and our use of a design with ample power. This is precisely the type of issue that has been shown to elicit a boomerang in other research (c.f., Nyhan & Reifler, 2010). We speculate about this briefly. Although people consistently counterargue attitude-discrepant information, attitude change is still possible if the evidence is sufficiently compelling (Ditto & Lopez, 1992; Ecker, Lewandowsky, Swire, & Chang, 2011). Perhaps the reason we failed to find a boomerang effect is simply that the fact-checking messages we used were more persuasive than those used in other studies, thereby making participants less likely to have their initial views bolstered through counterargument (Byrne & Hart, 2009). Alternatively, it may be that our ability to detect such an effect was limited by our direct measurement of belief in the disputed information. Research on the continued influence effect (CIE) has shown that *inferences* based on inaccurate information persist well after the information itself has been discredited (Seifert, 2002). Perhaps an inference-based measure of the inaccurate belief would have revealed a boomerang effect. This argument is made less compelling by the fact that our measurement approach is similar to that used by Nyhan and Reifler (2010), who were able to clearly and repeatedly demonstrate this

effect. Furthermore, researchers who have examined the boomerang effect within the CIE framework have not found evidence of it (Ecker et al., 2011). More research into the causes and conditions under which this effect occurs is needed.

Limitations

This study has a number of limitations that deserve mention. First, we considered a single misperception that happened to be most common among opponents of the Islamic cultural center, and who were disproportionately conservative Republicans. It is possible that a different dynamic would be found among project supporters, although the limited empirical evidence to date suggests that liberals are also inclined to interpret fact-checking information in a manner consistent with their own views (Nyhan & Reifler, 2010). It is also conceivable that responses would differ if other topics were employed, though there is nothing in the existing literature to suggest that these results are specific to the events studies. Nevertheless, replicating this study simultaneously across various policy contexts and media formats (e.g., newspaper, television) would provide further evidence about the role of contextual cues in fact checking and corrective effects.

Second, we do not measure the long-term effects of the correction. Other work has shown that the effects of corrections change over time (Schwarz et al., 2007), but this study does not allow us to say anything about how these effects play out. More research on this topic is merited.

A third consideration is that we omitted an experimental condition that included both the Imam's objectionable statement and the photograph showing him in more traditional Islamic dress. Our primary interest was in isolating two factors that may activate audience cognitions leading them to resist corrective efforts. We speculate that combining these factors would have significantly decreased the effectiveness of a corrective message, but it is unknown whether this would be sufficient to produce a boomerang effect. The failure to find a boomerang effect also raises another consideration. It is possible that a boomerang effect would have occurred among those who had a strong negative psychological reaction based on *intended construct activation* by the corrective messages. However, because we were focused on unintended construct activation, we did not measure psychological reactance and therefore are unable to test this possibility.

Finally, and most importantly, we acknowledge that we have not manipulated naïve theories directly; instead, we focused our efforts on crafting a theoretical explanation for a real-world phenomenon—namely, people's resistance to fact-checking messages—and then studying that phenomenon in a controlled setting. Thus, we have traded internal for ecological validity, and as a consequence we must acknowledge that other explanations are possible. For example, perhaps the patterns witnessed here are due to an affective response to the information provided, not to the naïve theories elicited by exposure. We may also have underestimated the role of the elaboration likelihood model of persuasion (Petty & Cacioppo, 1986) and the possibility that individuals engage in peripheral processing of fact-checking messages.

Without a more direct manipulation of activated naïve theories, we cannot know with certainty whether the mechanisms described here are accurate. Nevertheless, our evidence is consistent with our theory, and we hope that these results will spur more research on this topic.

Practical implications

These findings pose a profound challenge to any organization that would seek to fight public misperceptions and increase understanding of political facts. In addition to avoiding repetition of the misinformation itself (e.g., just providing the facts may be more effective than listing and denying a series of popular myths, see Schwarz et al., 2007), we must also think carefully about the nature of factually accurate information that we include, and we must find ways to defuse the harmful effects of inadvertently activating negative heuristics. The results of this study suggest the most straightforward method of maximizing the corrective effect of a fact-checking article is to avoid including information that activates stereotypes or generalizations—negative naïve theories—which make related cognitions more accessible and misperceptions more plausible. This strategy has a number of obvious limitations, though. There may be benefits to excluding factually accurate context, but this exclusion comes at a cost to the larger political conversation. It also implies that reference to any group that is perceived as threatening by those who hold the misperception, including racial, ethnic, or religious groups, can potentially undermine corrective effects. Whether you believe that including such information is important or not, it is likely to be impractical in most situations, for there is no simple way to systematically exclude all cues to these categories.

Fact checking is a critical function of the news media given the sometimes startling prevalence of political misperceptions in the U.S. today. Unfortunately, effectively correcting these misperceptions is quite difficult. In this paper we have identified a number of factors that counteract the benefits of providing factual information by triggering decision-making heuristics that, though effective in many cases, prove detrimental to individuals' ability to process novel political information. Our hope is that in offering these insights, we will stimulate further research in this critical area and, more pragmatically, that our suggestions may prove useful to news providers who must grapple every day with how best to present the factual information that they wish to convey to their audiences.

Acknowledgments

References

Anderson, C. A., Lepper, M. R. & Ross, L. (1980). Perseverance of social theories: The role of explanation in the persistence of discredited information. *Journal of Personality and Social Psychology*, **39**(6), 1037−1049. doi: 10.1037/h0077720

Anderson, C. A. & Lindsay, J. J. (1998). The development, perseverance, and change of naïve theories. *Social Cognition*, **16**, 8–30.

Bar, M., Neta, M. & Linz, H. (2006). Very first impressions. *Emotion*, **6**(2), 269–278. doi: 10.1037/1528-3542.6.2.269

Berinsky, A. J. (2012). The birther sare back. Accessed Feb. 3, 2012, at http://today.yougov.com/news/2012/02/03/birthers-are-back/

Brehm, J. (1966). *A theory of psychological reactance*. New York: Academic Press.

Byrne, S. & Hart, P. S. (2009). The 'boomerang' effect: A synthesis of findings and a preliminary theoretical framework. *Communication Yearbook*, **33**, 3–37.

Cohen, J., Cohen, P., West, S. G. & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences*. Mahwah, NJ: Lawrence Erlbaum Associates.

Coll, S. (2004). *Ghost wars: The secret history of the CIA, Afghanistan, and bin Laden, from the Soviet invasion to September 10, 2001*. New York, NY: Penguin Press.

DiFonzo, N. & Bordia, P. (2007). *Rumor psychology: Social and organizational approaches*. Washington, DC: American Psychological Association.

Ditto, P. H. & Lopez, D. F. (1992). Motivated skepticism. *Journal of Personality and Social Psychology*, **63**(4), 568–584.

Downs, A. (1957). *An economic theory of democracy*. New York, NY: Harper and Row.

Ecker, U., Lewandowsky, S., Swire, B. & Chang, D. (2011). Correcting false information in memory: Manipulating the strength of misinformation encoding and its retraction. *Psychonomic Bulletin & Review*, **18**(3), 570–578. doi: 10.3758/s13423-011-0065-1

Gibbs, W. W. (2005). Considerate computing. *Scientific American*, **292**(1), 54–61.

Gilens, M. (2001). Political ignorance and collective policy preferences. *American Political Science Review*, **95**(2), 379–396.

Hall, C. C., Goren, A., Chaiken, S. & Todorov, A. (2009). Shallow cues with deep effects: Trait judgments from faces and voting decisions. In E. Borgida, C. M. Federico, & J. L. Sullivan (Eds.), *The political psychology of democratic citizenship* (pp. 73–99). New York, NY: Oxford University Press.

Harrington, B. (2009). *Deception: From ancient empires to Internet dating*. Stanford, CA: Stanford University Press.

Hart, P. S. & Nisbet, E. C. (2012). Boomerang effects in science communication: How motivated reasoning and identity cues amplify opinion polarization about climate mitigation policies. *Communication Research*, **39**, 701–723. doi: 10.1177/0093650211416646

Higgins, E. T., Rholes, W. S. & Jones, C. R. (1977). Category accessibility and impression formation. *Journal of Experimental Social Psychology*, **13**(2), 141–154. doi: 10.1016/s0022-1031(77)80007-3

Jamieson, K. H. & Gottfried, J. A. (2010). Are there lessons for the future of news from the 2008 presidential campaign? *Daedalus*, **139**(2), 18–25.

Jones, N. A., Ross, H., Lynam, T., Perez, P. & Leitch, A. (2011). Mental models: An interdisciplinary synthesis of theory and methods. *Ecology and Society*, **16**(1), 46.

Kosloff, S., Greenberg, J., Schmader, T., Dechesne, M. & Weise, D. (2010). Smearing the opposition: Implicit and explicit stigmatization of the 2008 U.S. presidential candidates and the current U.S. president. *Journal of Experimental Psychology. General*, **139**(3), 383–398. doi: 10.1037/a0018809

Kuklinski, J. H., Quirk, P. J., Jerit, J., Schwieder, D., & Rich, R. F. (2000). Misinformation and the currency of democratic citizenship. *The Journal of Politics*, **62**(03), 790–816. DOI: 10.1111/0022-3816.00033.

Lau, R. R., Andersen, D. J. & Redlawsk, D. P. (2008). An exploration of correct voting in recent U.S. presidential elections. *American Journal of Political Science*, **52**(2), 395–411. doi: 10.1111/j.1540-5907.2008.00319.x

Lord, C. G., Lepper, M. R. & Preston, E. (1984). Considering the opposite: A corrective strategy for social judgment. *Journal of Personality and Social Psychology*, **47**(6), 1231–1243. doi: 10.1037/0022-3514.47.6.1231

Lord, C. G., Ross, L. & Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology*, **37**(11), 2098–2109.

Lupia, A. & McCubbins, M. D. (1998). *The democratic dilemma*. Cambridge: Cambridge University Press.

Macrae, C. N. & Bodenhausen, G. V. (2000). Social cognition: Thinking categorically about others. *Annual Review of Psychology*, **51**(1), 93–120. doi: 10.1146/annurev.psych.51.1.93

Meffert, M. F., Chung, S., Joiner, A. J., Waks, L., & Garst, J. (2006). The Effects of Negativity and Motivated Information Processing During a Political Campaign. *Journal of Communication*, **56**(1), 27–51. doi: 10.1111/j.1460-2466.2006.00003.x

Moore, M. (2004). *Fahrenheit 9/11*. Miramax Films.

Morrison, K. R. & Ybarra, O. (2009). Symbolic threat and social dominance among liberals and conservatives: SDO reflects conformity to political values. *European Journal of Social Psychology*, **39**(6), 1039–1052. doi: 10.1002/ejsp.606

Nacos, B. L. & Torres-Reyna, O. (2006). *Fueling our fears: Stereotyping, media coverage, and public opinion of Muslim Americans*. New York, NY: Rowman & Littlefield Publishers.

Newport, F. (2011, March 9). Republicans and Democrats disagree on Muslim hearings. Accessed March 9, 2011, at http://www.gallup.com/poll/146540/republicans-democrats-disagree-muslim-hearings.aspx

Nisbet, E. C., Ostman, R. & Shanahan, J. (2009). Public opinion toward Muslim Americans: Civil liberties and the role of religiosity, ideology, and media use. In A. H. Sinno (Ed.), *Muslims in Western politics* (pp. 161–199). Bloomington, IN: Indiana University Press.

Nyhan, B. & Reifler, J. (2010). When corrections fail: The persistence of political misperceptions. *Political Behavior*, **32**(2), 303–330. doi: 10.1007/s11109-010-9112-2

Petty, R. E. & Cacioppo, J. T. (1986). *Communication and persuasion: Central and peripheral routes to attitude change*. New York, NY: Springer-Verlag.

Petty, R. E. & Wegener, D. T. (1998). Attitude change: Multiple roles for persuasion variables. In D. Gilbert, S. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., pp. 323–390). New York, NY: McGraw-Hill.

Pew Forum on Religion & Public Life (December, 2012) The Global Religious Landscape. Accessed 12/18/2012 at http://www.pewforum.org/global-religious-landscape.aspx

Pintak, L. (2006). *Reflections in a bloodshot lens: America, Islam and the war of ideas*. Ann Arbor, MI: Pluto.

Popkin, S. L. (1991). *The reasoning voter: Communication and persuasion in presidential campaigns*. Chicago, IL: University of Chicago Press.

Pratto, F., Sidanius, J., Stallworth, L. M. & Malle, B. F. (1994). *Social dominance orientation: A personality variable predicting social and political attitudes* **67**. Washington, DC: American Psychological Association.

Redlawsk, D. P. (2002). Hot Cognition or Cool Consideration? Testing the Effects of Motivated Reasoning on Political Decision Making. *Journal of Politics*, **64**(4), 1021−1044.

Sanna, L. J. & Schwarz, N. (2003). Debiasing the hindsight bias: The role of accessibility experiences and (mis)attributions. *Journal of Experimental Social Psychology*, **39**(3), 287−295. doi: 10.1016/s0022-1031(02)00528-0

Schwarz, N., Sanna, L. J., Skurnik, I. & Yoon, C. (2007). Metacognitive experiences and the intricacies of setting people straight: Implications for debiasing and public information campaigns. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (**39**, pp. 127−161). New York, NY: Academic Press.

Seib, P. M. (2005). The news media and ''the clash of civilizations''. In P. M. Seib (Ed.), *Media and conflict in the twenty-first century*. New York, NY: Palgrave Macmillan.

Seifert, C. M. (2002). The continued influence of misinformation in memory: What makes a correction effective?. In H. R. Brian (Ed.), *Psychology of learning and motivation* (**41**, pp. 265−292). New York, NY: Academic Press.

Shaheen, J. G. (2001). *Reel bad Arabs: How Hollywood vilifies a people*. New York, NY: Olive Branch Press.

Shtulman, A. & Valcarcel, J. (2012). Scientific knowledge suppresses but does not supplant earlier intuitions. *Cognition*, **124**(2), 209−215. doi: 10.1016/j.cognition.2012.04.005

Sidanius, J. & Pratto, F. (1999). *Social dominance: An intergroup theory of social hierarchy and oppression*. New York, NY: Cambridge University Press.

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

Thomsen, L., Green, E. G. T. & Sidanius, J. (2008). We will hunt them down: How social dominance orientation and right-wing authoritarianism fuel ethnic persecution of immigrants in fundamentally different ways. *Journal of Experimental Social Psychology*, **44**(6), 1455−1464. doi: 10.1016/j.jesp.2008.06.011

Tversky, A. & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, **5**(2), 207−232. doi: 10.1016/0010-0285(73)90033-9

World Public Opinion (2006). *Percentage of Americans believing Iraq had WMD rises*. Washington, DC: World Public Opinion: Global Public Opinion on International Affairs.

Appendix: Experimental stimuli

Rebuttal text

There were two versions of the rebuttal text, both of which are summarized below. The low-controversy rebuttal consisted of the unbolded text. The high-controversy version was created by dropping the underlined text and adding the bolded text:

Factcheck.org, an award-winning nonpartisan, nonprofit ''consumer advocate'' for voters that aims to reduce the level of deception and confusion in U.S. politics, examined the issue. They wrote that they found no evidence that Feisal Abdul Rauf was anti-American or a terrorist-sympathizer. In fact, Mr. Abdul Rauf has been a vocal critic of Islamic extremists and has condemned their use of violent attacks. He has a history of supporting U.S. initiatives designed to stop Muslim extremism that dates back to the Bush administration. **Although Mr. Adbul Rauf has said that**

the U.S. bears responsibility for harm caused by its policies toward the Middle East, he is quick to point out that terrorism is never justified.

In both 2007 and 2010 he traveled through the Middle East to talk about religious tolerance as part of a speaker program sponsored by the American State Department. And in a recent interview with the *Washington Post* he spoke about the group proposing to build a Muslim community center near Ground Zero: "We are not the extremists. We are that vast majority of Muslims who stand up against extremism and provide a voice in response to the radical rhetoric. **Our mission is to interweave America's Muslim population into mainstream society. We are a Muslim-American force for promoting the universal values of justice and peaceful coexistence**..."

Rebuttal images

The low-controversy rebuttal was paired with images portraying Imam Rauf in either (a) Western-style attire or (b) Middle-Eastern-style attire. The study used color images, which are available from the first author upon request.

(a)

(b)