

ISSA Proceedings 2014 - The Failure Of Fact-Checking

Abstract: Fact-checking rests on a foundation that is desirable: an educated citizenry, informed of the facts, will make a rational decision. Unfortunately, the theory of motivated reasoning suggests prior attitudes strongly influence the process. This paper reports the results from two studies ($n=456$) that investigated the effectiveness of fact-checking in the context of ObamaCare. The results of the studies confirm the real problem for fact-checking: prior attitudes intervened to reduce the utility of the fact check.

Keywords: fact-checking, motivated reasoning, Obama, political affiliation, Romney

1. Introduction

In an effort to combat a new wave of false and misleading political advertisements, American journalists in the early 1990's shed their tendency merely to report politicians' claims and instead took up the challenge to report their truthfulness. These new adwatches were meant provide the public with the information necessary to make an informed decision. Journalists embraced their role as the arbiters of truth with the hope "that prospective voters will use information about misleading ads to discount their claims and turn away from candidates who ads lack veracity" (Frantzich, 2002, p. 35). The idea that voters would rely on evidence and rationally choose a candidate is an ideal that is firmly rooted in democracy. But, what is less clear is whether voters use the information provided in adwatches to make a "good" decision.

Early research investigated the effectiveness of television news adwatches and found mixed results. Some studies found that adwatches were not effective. Ansolabehere and Iyengar (1996) found adwatches backfired since "the candidate who was scrutinized by the media enjoyed increased support among those who watched an ad-watch report" (p. 82). Pfau and Loudon (1994) report mixed results: one candidate gained support while the opponent in the race faced lower support. Jamieson and Cappella (1997) criticized both studies over methodological problems. The result found by Ansolabehere and Iyengar, they argued, was the result of a generally favorable fact-check conclusion that

“supported the gist of the claims made in the ad” (p. 16). Similarly, the conclusion found by Pfau and Louden was the result of “inviting comparison between candidates” (p. 14) since they were in the same race. In contrast to these studies, Cappella and Jamieson (1994) found adwatches were effective. “The adwatches appear to do precisely what they are designed to accomplish, namely put the claims of the ad in context so that the ad is judged less fair and less important” (Cappella & Jamieson, 1994, p. 355). Other studies also have found exposure to fact-checking to be effective (Garrett, Nisbet, & Lynch, 2013; Gottfried, Hardy, Winneg, & Jamieson, 2012; O’Sullivan & Geiger, 1995). Contemporary research in political science and social psychology has investigated the topic under the banner of “corrections” to political misinformation also with mixed results. Several studies have documented the failure of new information to correct the misinformation (Bullock, 2006; Kuklinski, Quirk, Jerit, Schwieder, & Rich, 2000; Nyhan & Reifler, 2010). In contrast, some research has found that corrections were effective at changing opinion (Gilens, 2001; Howell & West, 2009; Kuklinski et al, 2000). In those cases, however, the public changed their mind because the information was so overwhelming that it “hit them between the eyes” (Kuklinski et al., 2000, p. 805).

How can we make sense of these inconclusive results? There is hope to find an answer. Referred to variously as motivated reasoning (Mercier & Sperber, 2011), the prior attitude effect (Taber & Lodge, 2006), belief perseverance (Bullock, 2006), biased assimilation (Lord, Ross, & Lepper, 1979), or an attitude congruency bias (Taber, Cann, & Kucsova, 2009), the theory is simple: individuals will “judge confirming evidence as relevant and reliable but disconfirming evidence as irrelevant and unreliable” and will “accept confirming evidence at face value while scrutinizing disconfirming evidence hypercritically” (Lord, Ross, & Lepper, 1979, p. 2099). In other words, people are more likely to believe what they already believe and are less likely to believe what they already reject. Political affiliation serves as an important source of bias in the interpretation of political information (Allen, Stevens, & Sullivan, 2009; Bullock, 2006; Gaines, Kuklinski, Quirk, Peyton, & Verkuilen, 2007; Nyhan & Reifler, 2010). Prior research on adwatches and fact-checking rarely accounted for the role of motivated reasoning. Some articles mention descriptive statistics for political affiliation, but most studies did not analyze the role of motivated reasoning, expressed most clearly as political affiliation and a partisan bias for the source, in the evaluation of the advertisement (Ansolabehere & Iyengar, 1996; Cappella &

Jamieson, 1995; Jamieson, 1992; McKinnon & Kaid, 1999; Milburn & Brown, 1995; Pfau & Loudon, 1994; O'Sullivan & Geiger, 1995). This is a major shortcoming in prior research. The influence of prior attitude is significant, and political affiliation is an important part of that attitude. This project attempts to remedy this shortcoming by explicitly addressing the role of prior attitude, especially political commitments, when evaluating statements subject to fact-check criticism. As such, it is guided by the following hypotheses and research question:

H1: Evaluation of the strength of the initial argument will be influenced by initial evaluation of the source and political affiliation, such that attitude congruent arguments will be rated higher and attitude incongruent arguments will be rated lower.

H2: Evaluation of the strength of the fact-check analysis will be influenced by initial evaluation of the source and political affiliation, such that attitude congruent arguments will be rated higher and attitude incongruent arguments will be rated lower.

H3: Final evaluation of the strength of the argument will be influenced by initial evaluation of the source and political affiliation, such that attitude congruent arguments will be rated higher and attitude incongruent arguments will be rated lower.

RQ1: Do other factors such as age, sex, level of education, or level of political interest create prior attitude effects when evaluating political messages or fact-check responses?

2. Methodology

Two studies were conducted in the fall of 2012 to investigate the topic. Both studies followed the same basic procedure and will be explained together. Experiment 1 focused on reactions to a statement by President Obama on the cost effectiveness of preventative health care. Experiment 2 focused on reactions to a statement by Mitt Romney on the cost savings associated with repealing Obamacare.

Participants initially completed a series of basic demographic questions, including age, sex, level of education, political affiliation and level of political interest. Participants also completed a feeling thermometer to express their attitude

toward Obama (Experiment 1) or Romney (Experiment 2). The main experiment involved three stages. First, respondents were shown a brief statement by either Obama (71 words) or Romney (62 words) and recorded their evaluation of the strength of the argument using a semantic differential scale (explained below). Then, respondents were shown a lengthy refutation of the argument (439 words in Experiment 1, 332 words in Experiment 2) by a fact-checking organization and recorded their evaluation of the strength of the fact-check statement using the same scale. Finally, participants were asked to re-evaluate the strength of the original claim using the same scale.

In both experiments, the fact-check analysis provided strong refutation of the original statement and definitively suggested the claim was false. The evidence in support of the fact-check conclusion came from politically neutral sources. In Experiment 1, the sources included the Congressional Budget Office, a report in the *New England Journal of Medicine*, and a study sponsored by the American Diabetes Association, American Heart Association, and the American Cancer Society. In Experiment 2, the sources included the Congressional Budget Office and detailed the revenue enhancements contained in Affordable Care Act.

Argument strength was measured using a semantic differential scale adapted from La France and Boster (2001). The pairs included correct-incorrect, valuable-not valuable, unsound-sound, poorly reasoned-well reasoned, and reasonable-unreasonable. Items were recoded so the negative element received the lowest score. Items were summed to create an overall evaluation for the argument. Scores ranged from 6 to 42. A low total score suggested the argument was weak while a high total score suggested the argument was strong. The scales were reliable (alpha reported for each experiment).

Attitude toward Obama and Romney was measured before and after the experiment using a feeling thermometer (0-100) typical in political research (ANES, 2008). Level of political involvement was measured with a single item that asked, "How interested are you in information about what's going on in government and politics?" (ANES, 2008). Respondents could select extremely interested, very interested, moderately interested, slightly interested, or not interested at all. Level of education was measured on a five-point ordinal scale (ANES, 2008). Respondents could select no high school diploma, high school diploma, some college (but no Bachelor's degree), Bachelor's degree, or education beyond a Bachelor's degree.

3. *Experiment 1*

Participants ($N=187$) were recruited from several communication classes (in exchange for course credit/extra-credit) and from Amazon's Mechanical Turk service (receiving payment ranging from \$.8 to \$1.5 for completing the survey). Participants ranged in age from 18 to 66 years of age ($M=29.12$, $SD=11.77$). A slight majority was female ($n=95$). Most participants had some education beyond a high school diploma ($n=149$). A third of the participants (33.7%) reported either an extremely high or very high interest in information about government and politics. About a quarter of the participants (26.7%) reported only some interest or no interest in similar information. Over a third of participants reported a political affiliation consistent with the Democratic Party (37.4%) compared with nearly a quarter that identified as Republicans (23.5%). Just more than a quarter identified as Independents (28.3%) with the remaining selecting some other political affiliation (10.7%).

The semantic differential scales used in the survey were highly reliable. Cronbach's alpha for the three scales were: initial evaluation of Obama's argument ($\alpha=.94$), evaluation of fact-check analysis ($\alpha=.92$), and re-evaluation of Obama's argument ($\alpha=.95$).

The first hypothesis suggested that people would be influenced by their prior attitudes when judging the initial strength of the argument made by President Obama. There was overwhelming support for the hypothesis. A one-way ANOVA of political affiliation was conducted on initial evaluation of the strength of the statement by President Obama. This analysis produced a statistically significant result, $F(3,183)=28.54$, $p<.001$, $\eta^2=.32$). Post hoc tests using the Bonferroni correction revealed a significant difference between Democrats ($M=35.8$, $SD=6.98$), who rated the argument strongest, and Republicans ($M=21.98$, $SD=8.30$), who rated the argument weakest, as well as Independents ($M=30.94$, $SD=8.01$) and others ($M=29.95$, $SD=8.63$). In fact, the only comparison that was not statistically significant was between Independents and others. In addition, correlation was used to test the relationship between attitude toward Obama and the initial evaluation of the argument. There was a strong association between the pretest thermometer rating for Obama and the initial evaluation of the argument, $r(185)=.64$, $p<.001$, $r^2=.41$. As ratings for Obama increased, so too did the evaluation of his argument.

The second hypothesis suggested that people would be influenced by their prior

attitudes when judging the strength of the fact-check analysis. There was support for this hypothesis. A one-way ANOVA of political affiliation was conducted on the evaluation of the strength of the fact-check analysis. This analysis produced a statistically significant result, $F(3,183)=4.76$, $p=.003$, $\eta^2=.07$. Post hoc tests using the Bonferroni correction revealed a significant difference between Democrats ($M=26.29$, $SD=8.01$), who rated the argument weakest, and Republicans ($M=31.66$, $SD=7.44$), who rated the argument strongest. There were no other statistically significant differences. In addition, correlation was used to test the relationship between attitude toward Obama and the evaluation of the fact-check analysis. There was a moderate negative correlation between the pretest thermometer rating for Obama and the evaluation of the fact-check analysis, $r(185)=-.33$, $p<.001$, $r^2=.11$. As ratings for Obama increased, evaluation of the fact-check criticism went down.

The third hypothesis suggested that people would continue to be influenced by their prior attitudes when making the final evaluation of the argument. A one-way ANOVA of political affiliation was conducted on the final evaluation of the strength of the statement by Obama. The analysis produced a statistically significant result, $F(3,183)=32.39$, $p<.001$, $\eta^2=.35$. Post hoc tests using the Bonferroni correction revealed a significant difference between Democrats ($M=32.41$, $SD=7.65$), who continued to rate the statement strongest, and Republicans ($M=17.41$, $SD=7.20$), who continued to rate the statement weakest, as well as Independents ($M=26.00$, $SD=8.86$). The only comparisons that were not significant were between and others ($M=27.10$, $SD=7.81$) and both Democrats and Independents. In addition, correlation was used to test the relationship between attitude toward Obama and the final evaluation of the statement. Even after exposure to a fact-check criticism, a stronger positive correlation was found between the pretest thermometer rating for Obama and the final evaluation of the statement, $r(185)=.67$, $p<.001$, $r^2=.45$.

Finally, to investigate the research question, a series of one-way ANOVAs were conducted to compare sex, level of education, and level of political interest with the initial evaluation of Obama's argument, the evaluation of the fact-check analysis, and the final evaluation of Obama's argument. In addition, given the importance of political affiliation, a series of two-way ANOVAs were conducted using sex, level of education, and level of political interest along with political affiliation on each evaluation. Correlation was used to compare age and the three

evaluations. Factors beyond political affiliation played almost no role in the evaluation of Obama's statement or the fact-check criticism.

A one-way ANOVA was conducted between sex and the three argument evaluations. No group differences based on sex were found for initial evaluation, $F(1,185)=.74$, $p=.39$, final evaluation, $F(1,185)=.09$, $p=.77$, or evaluation of the fact-check analysis, $F(1,185)=.312$, $p=.577$. A two-way ANOVA was conducted between sex and political affiliation on the three argument evaluations. There was no main effect for sex and no interaction effect between political affiliation and sex for any of the three evaluations.

A one-way ANOVA found no group differences based on level of education for initial evaluation, $F(4,182)=.805$, $p=.524$, final evaluation, $F(4,182)=.381$, $p=.882$, or evaluation of the fact-check analysis, $F(4,182)=.875$, $p=.480$. In addition, a two-way ANOVA was conducted between level of education and political affiliation on the three argument evaluations. There was no main effect for level of education and no interaction effect political affiliation and level of education for any of the three evaluations.

Differences in political interest were found for initial evaluation, $F(4,182)=2.582$, $p=.039$. No group differences based on political interest were found for the evaluation of the fact-check analysis, $F(4,182)=1.701$, $p=.152$, or for the final evaluation, $F(4,182)=1.467$, $p=.214$. A two-way ANOVA was conducted between level of interest and political affiliation on the three argument evaluations. There was no main effect for level of political interest and no interaction effect between political affiliation and level of political interest for any of the three evaluations. In general, sex and level of education played no role in the evaluation of the competing political statements. Level of political interest played a very small role in the assessment of the statements. Any role played by political interest was dwarfed by political affiliation.

Correlation was used to test the relationship between age and the three argument evaluations. Age was not associated with the initial evaluation of the argument, $r(185)=.11$, $p=.14$, the final evaluation of the argument, $r(185)=.02$, $p=.77$, or the evaluation of the fact-check criticism, $r(185)=.02$, $p=.84$

4. *Experiment 2*

The results from the first experiment were conclusive in favor of a motivational

bias in the processing of fact-check information. But, for balance, a second experiment was conducted using Mitt Romney as the source and an anti-Obamacare argument as the material for evaluation. Participants ($N=269$) were recruited from introductory communication classes (in exchange for course credit/extra-credit) and from Amazon's Mechanical Turk service (receiving payment of \$.8 for completing of the survey). Participants ranged from 18 to 69 years of age ($M=28.92$, $SD=11.19$). A slight majority was male ($n=138$). Most participants had an education beyond a high school diploma (82.9%). Nearly half of the participants (45.4%) reported either an extremely high or very high interest in information about government and politics. Only 18.6% reported only some interest or no interest in similar information. A plurality of participants reported a political affiliation consistent with the Democratic Party (38.1%). Republicans (26.7%) and Independents (26.3%) were the next most common affiliation. A few participants (8.5%) reported some other political affiliation.

The semantic differential scales used in the survey were reliable. Cronbach's alpha for the three scales were: initial evaluation of Romney's argument ($\alpha=.96$), evaluation of fact-check analysis ($\alpha=.96$), and re-evaluation of Mitt Romney's argument ($\alpha=.97$).

The first hypothesis suggested that people would be influenced by their prior attitudes when judging the initial strength of the argument made by Mitt Romney. There was overwhelming support for the hypothesis. A one-way ANOVA of political affiliation was conducted on initial evaluation of the strength of the statement by Romney. This analysis produced a statistically significant result, $F(3,265)=60.78$, $p<.001$, $\eta^2=.41$. Post hoc tests using the Bonferroni correction revealed significant differences between Republicans ($M=33.18$, $SD=7.82$), who rated the argument strongest, and Democrats ($M=15.18$, $SD=8.17$), who rated the argument weakest, as well as both Independents ($M=20.93$, $SD=9.76$) and others ($M=24.74$, $SD=10.73$). In fact, all other comparisons were statistically significant except between Independents and others. In addition, correlation was used to compare prior attitude toward Romney and the initial evaluation of his argument. There was a strong, positive correlation between pretest thermometer rating for Romney and the initial evaluation of his statement, $r(267)=.78$, $p<.001$, $r^2=.61$. Increasing evaluations for Romney were associated with increasing evaluations for his statement.

The second hypothesis suggested that people would be influenced by their prior

attitudes when judging the strength of the argument presented in the fact-check analysis. A one-way ANOVA of political affiliation was conducted on evaluation of the fact-check. The analysis produced a statistically significant result, $F(3,265)=14.83$, $p<.001$, $\eta^2=.14$. Post hoc tests using the Bonferroni correction revealed significant differences between Republicans ($M=23.42$, $SD=9.21$), who rated the argument weakest, and Democrats ($M=32.14$, $SD=10.80$), Independents ($M=32.28$, $SD=7.67$) and others ($M=30.13$, $SD=7.91$). There were no other statistically significant differences. In addition, correlation was used to compare prior attitude toward Romney and the evaluation of the fact-check criticism. There was a moderate, negative correlation between pretest thermometer rating for Romney and the evaluation of the fact-check criticism, $r(267)=-.46$, $p<.001$, $r^2=.21$. Increasing evaluation of Romney was associated with a decreasing evaluation of the fact-check criticism.

The final hypothesis suggested that people would continue to be influenced by their prior attitudes, even after exposure to a fact-check criticism, when evaluating the strength of Romney's argument. A one-way ANOVA of political affiliation was conducted on the final evaluation of the statement made by Romney. The analysis produced a statistically significant result, $F(3,265)=81.70$, $p<.001$, $\eta^2=.48$. Post hoc tests using the Bonferroni correction revealed significant differences between Republicans ($M=30.04$, $SD=8.03$), who continue to rate the argument strongest, and Democrats ($M=11.26$, $SD=5.97$), Independents ($M=16.14$, $SD=9.62$) and others ($M=20.35$, $SD=9.59$). In fact, all of the comparisons were statistically significant except between Independents and others. In addition, correlation was used to compare prior attitude toward Romney and the final evaluation of his statement. There was an even stronger, positive correlation between pretest thermometer rating for Romney and the final evaluation of his statement, $r=.79$, $p<.001$, $r^2=.63$.

Finally, the research question was investigated using a series of one-way ANOVAs between sex, level of education, and level of political interest with all three argument evaluations. In addition, a series of two-way ANOVAs compared political affiliation along with sex, level of education, and level of political interest with all three argument evaluations. Correlation also was used to identify a relationship between age and the three argument evaluations. Very few factors had a meaningful impact on argument evaluation.

There were no group differences between sex and the initial evaluation of the

argument, $F(1,267)=.004$, $p=.95$, evaluation of the fact-check, $F(1,267)=2.24$, $p=.14$, or the final evaluation of the argument, $F(1,267)=1.25$, $p=.27$. In addition, a two-way ANOVA with sex and political affiliation found no main effect for sex and no interaction effect with political affiliation for the initial evaluation of the argument, the final evaluation of the argument, or for the evaluation of the fact-check criticism.

There were no group differences between level of education and the initial evaluation of the argument, $F(3,265)=2.27$, $p=.08$, final evaluation of the argument, $F(3,265)=1.09$, $p=.36$, or for the evaluation of the fact-check criticism, $F(3,265)=.23$, $p=.88$. In addition, a two-way ANOVA with level of education and political affiliation found no main effect for level of education and no interaction effect with political affiliation for the initial evaluation of the argument, the final evaluation of the argument, or for the evaluation of the fact-check criticism.

There were no group differences between level of political interest and the initial evaluation of the argument, $F(4,264)=1.61$, $p=.17$, final evaluation of the argument, $F(4,264)=0.83$, $p=.51$, or for the evaluation of the fact-check criticism, $F(4,264)=1.16$, $p=.33$. A two-way ANOVA with level of political interest and political affiliation found a main effect for both political affiliation, $F(3,250)=42.12$, $p<.001$, $\eta^2=.34$, and level of political interest, $F(4,250)=4.49$, $p=.002$, $\eta^2=.07$, but not for the interaction. Post hoc tests using the Bonferroni correction found significant differences between those with no interest at all, who rated the argument lowest, and those with a slight interest and those with a moderate interest. No other levels of political interest were significantly different. Similarly, a two-way ANOVA found a main effect for both political affiliation, $F(3,250)=49.60$, $p<.001$, $\eta^2=.37$, and level of political interest, $F(4,250)=2.67$, $p=.03$, $\eta^2=.04$. Post hoc tests using the Bonferroni correction found no significant differences between any of the groups. A two-way ANOVA with level of political interest and political affiliation found no main effect and no interaction effect with political affiliation for the evaluation of the fact-check criticism.

Correlation was used to test the relationship between age and the three argument evaluations. Age was not associated with the initial evaluation of the argument, $r(267)=-.094$, $p=.12$, the final evaluation of the argument, $r(267)=-.07$, $p=.24$, or the evaluation of the fact-check criticism, $r(267)=.02$, $p=.74$

5. Conclusion

Fact-checking represents a laudable goal in a democracy. In an effort to help shape public opinion, fact-checking provides citizens with the “facts” necessary to evaluate competing political claims. The public comes to any political dispute with some prior opinion. But, once exposed to a fact-check, they *should* update their opinion on the basis of the new information. Fact-checking, then, *should* promote a broad consensus on the topic. When the scrutinized claims generally are true, the public can have confidence in their prior opinion. But, when the scrutinized claims generally are false, exposure to a fact-check analysis ought to undermine the prior opinion and serve as the justification for a new opinion.

Unfortunately, the results from both experiments reported in this paper suggest that reality is far from the ideal. In both cases, prior partisan attitudes strongly influenced the evaluation of the arguments. The initial evaluation of the argument was shaped by prior commitments. Prior attitude toward the source and political affiliation were strong predictors of the initial evaluation of the statement: Proattitudinal messages were supported and counterattitudinal messages were rejected. The same commitments shaped reaction to the fact-check analysis. The criticism levelled by the fact-check was strong, well-supported from neutral sources, and unambiguously concluded the initial claim was false. Yet, in both studies, prior commitments were strong predictors of the evaluation of the fact-check claims: attitude congruent messages were rated much higher than attitude incongruent messages. Finally, even after exposure to a strong fact-check analysis, prior attitudes continued to influence the evaluation of the political statement: attitude consistent messages continued to be supported and attitude inconsistent messages continued to be rejected.

It could be argued that the results of these experiments confirm that fact-checking is effective. After all, the evaluations of the statements were lower after exposure to the fact-check criticism. But, such a conclusion is not warranted for two reasons. First, the evaluations did not decline substantially. In both experiments, those in the proattitudinal conditional (Democrats in Experiment 1 and Republicans in Experiment 2) maintained evaluations that were very positive. While partisans reduced their evaluations, the average scores were still very positive. Exposure to a strongly worded criticism that seriously challenged the validity of the claim resulted in only a minor adjustment in evaluation by committed partisans. With average scores still above 30 (on a scale to 42), it is clear that partisans continue to believe the claim even in the face of a relentless

challenge. Second, fact-checking ought to cause a convergence of opinion. The ideal of fact-checking is premised on the idea that there is a “truth” and that exposing the public to the “correct” information will cause them to reject the misleading statements. Bayesian updating suggests that fact-checking should result in a convergence of opinion (Bartels, 2002, p. 122). “Two groups of Bayesian learners exposed to the same set of information should inexorably come to see the world in the same way” (Grynaviski, 2006, p. 331). Unfortunately, fact-checking does not encourage this effect. Instead of a convergence of opinion, the minor updating by committed partisans prevents agreement. The lack of agreement on the facts, confirmed by the continuing gap in the final evaluation of the statement, suggests that partisanship continues to influence the evaluation of the statement even after exposure to a strong fact-check criticism.

Motivated reasoning provides a useful explanation for these results. Competing political claims, and the statements of fact-checkers, invoke partisan interests that prevent a rational assessment of the information. As a result, messages that are consistent with prior commitments are viewed as strong and not subject to scrutiny while messages that are inconsistent are viewed as weak and actively scrutinized.

It would be easy to interpret these results as a call to abandon fact-checking. That is not the intent of this project. Rather, this is a call to encourage those engaged in fact-checking to move beyond basic descriptions of true and false. Merely identifying one side as making a false claim is not likely sufficient to change public opinion. Where Jamieson (1992) advanced the utility of adwatches by providing a visual “grammar” to make them more effective, this research highlights the importance of creating a “motivational” grammar to make them more compelling. Fact-checking should not be abandoned. But, fact-checkers need to be cognizant that counterattitudinal messages will be actively scrutinized and they must develop strategies to make their messages more compelling in the face of strong pressures to reject their claims. Merely hoping that the audience will be motivated by accuracy goals (Taber & Lodge, 2006) will not be sufficient.

Several limitations of this project deserve mention. First, study participants mainly were drawn from Amazon’s Mechanical Turk online labor market. While the service has been shown to be a reliable source for participants for academic research (Buhrmester, Kwang, & Gosling, 2011; Horton, Rand, & Zeckhauser, 2011; Mason & Suri, 2012), future research should attempt to gather responses

from a random sample. Moving from a convenience sample to a random sample would provide more confidence in the generalizability of the conclusions.

Second, this study addressed a single issue with a single, brief exposure to a fact-check criticism. While it appears that exposure to a single fact-check was not effective at changing opinion, it is possible that fact-checking could serve to change opinion in the long-term. Future studies should attempt to study the effect over many months, especially with repeated exposure to the criticism. In addition, fact-checking could serve to induce media coverage of misleading information, prompting repetition of the weakness of the claim. Repeated exposure, over a significant period of time, could provide the basis for opinion change, even if gradual and small. Additional research is needed to test the long-term effects of fact-checking on public opinion.

References

- Allen, B., Stevens, D., & Sullivan, J. (2009). Not the whole story: From punditry to partisan processing of the power of the truth test. Paper presented at 2009 American Political Science Association Annual Meeting, Toronto, Canada. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1449892.
- American National Election Study. (2008). American National Election Study, 2008: Pre- and post-election survey: Data collection instrument. ICPSR 25383. Retrieved from <http://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/25383>.
- Ansolabehere, S., & Iyengar, S. (1996). Can the press monitor campaign advertising? An experimental study. *Harvard International Journal of Press/Politics*, 1, 72-86.
- Bartels, L. M. (2002). Beyond the running tally: Partisan bias in political perceptions. *Political Behavior*, 24, 117-150.
- Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's mechanical turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science*, 6, 3-5.
- Bullock, J. G. (2006, October). Partisanship and the enduring effects of false political information. Retrieved from http://www.nyu.edu/gsas/dept/politics/seminars/bullock_f06.pdf.
- Cappella, J. N., & Jamieson, K. H. (1994). Broadcast adwatch effects. *Communication Research*, 21, 342-365.
- Frantzich, S. (2002). Watching the watchers: The nature and content of campaign ad watches. *Harvard International Journal of Press/Politics*, 7, 34-57.

- Gaines, B. J., Kuklinski, J. H., Quirk, P. J., Peyton, B., & Verkuilen, J. (2007). Same facts, different interpretations: Partisan motivation and opinion on Iraq. *Journal of Politics*, 69, 957-974.
- Garrett, R. K., Nisbet, E. C., & Lynch, E. K. (2013). Undermining the corrective effects of media-based political fact checking? The role of contextual cues and naïve theory. *Journal of Communication*, 63, 617-637.
- Gilens, M. (2001). Political ignorance and collective policy preferences. *American Political Science Review*, 95, 379-396.
- Gottfried, J. A., Hardy, B. W., Winneg, K. M., & Jamieson, K. H. (2013). Did fact checking matter in the 2012 presidential campaign? *American Behavioral Scientist*, 57, 1558-1567.
- Grynaviski, J. D. (2006). A Bayesian learning model with applications to party identification. *Journal of Theoretical Politics*, 18, 323-346.
- Horton, J. J., Rand, D. G., & Zeckhauser, R. J. (2011). The online laboratory: Conducting experiments in a real labor market. *Experimental Economics*, 14, 399-425.
- Howell, W. G., & West, M. R. (2009). Educating the public. *Education Next*, 9(3), 41-47.
- Jamieson, K. H. (1992). *Dirty politics: Deception, distraction, democracy*. New York: Oxford University Press.
- Jamieson, K. H., & Cappella, J. N. (1997). Setting the record straight: Do ad watches help or hurt? *Harvard International Journal of Press/Politics*, 2, 13-22.
- Kuklinski, J. H., Quirk, P. J., Jerit, J., Schwieder, D., & Rich, R. F. (2000). Misinformation and the currency of democratic citizenship. *Journal of Politics*, 62, 790-816.
- La France, B. H., & Boster, F. J. (2001). To match or mismatch? That is only one important question. *Communication Monographs*, 68, 211-234.
- 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, 2098-2109.
- Mason, W., & Suri, S. (2012). Conducting behavioral research on Amazon's mechanical turk. *Behavior Research Methods*, 44, 1-23.
- McKinnon, L. M., & Kaid, L. L. (1999). Exposing negative campaigning or enhancing advertising effects: An experimental study of adwatch effects on voters' evaluations of candidates and their ads. *Journal of Applied Communication*, 27, 217-236.
- Mercier, H., & Sperber, D. (2011). Why do humans reason? Arguments for an

argumentative theory. *Behavioral and Brain Sciences*, 34, 57-111.

Milburn, M. A., & Brown, J. (1995). Busted by the ad police: Journalists' coverage of political campaign ads in the 1992 presidential campaign. *Research Paper R-15*.

Nyhan, B. & Reifler, J. (2010). When corrections fail: The persistence of political misperceptions. *Political Behavior*, 32, 303-330.

O'Sullivan, P. B., & Geiger, S. (1995). Does the watchdog bite? Newspaper ad watch articles and political attack ads. *Journalism and Mass Communication Quarterly*, 72, 771-785.

Pfau, M., & Loudon, A. (1994). Effectiveness of adwatch formats in deflecting political attack ads. *Communication Research*, 21, 325-341.

Taber, C. S., Cann, D., & Kucsova, S. (2009). The motivated processing of political arguments. *Political Behavior*, 31, 137-155.

Taber, C. S., & Lodge, M. (2006). Skepticism in the evaluation of political beliefs. *American Journal of Political Science*, 50, 755-769.