Debunked, but not defunct: Continued influence of corrected negative misinformation

Dian van Huijstee, Ivar Vermeulen, and Peter Kerkhof

VU University Amsterdam

In modern mediated society, much of the information we are exposed to is untrue. One way to reduce harmful effects of such misinformation (e.g., for organizations, or public health) is by fact-checking—the process of evaluating the accuracy of statements, and correcting these if necessary (Graves, 2013). However, being consciously aware that information is corrected (and thus, in fact, is misinformation) does not always neutralize its influence (Lewandowsky et al., 2012). This phenomenon is called *continued influence*, indicating a distinction between factual belief in misinformation and its persuasive effects.

Prior studies have suggested an asymmetry in continued influence: whereas corrections of positive information may reverse initial persuasive effects (Nyhan & Reifler, 2010), persuasive effects of negative information seem more resistant to corrections (Ecker et al., 2019)—possibly because negative information is often judged more credible than positive information (Pan & Chiou, 2011) and thus harder to discredit. Also, negative information is perceived as more "diagnostic" for selection behavior (Baumeister et al., 2001) and therefore harder to disregard.

This suggested asymmetry in continued influence has never been investigated within a single study. We present two studies filling this gap by testing whether indeed continued influence effects of debunked negative information exceed those of debunked positive

information. Both use a response-latency based measure for attitudes; hypotheses were pre-registered at AsPredicted[1].

Study 1

Method. 387 MTurk participants evaluated images of hotels in a reaction-time task (positive/negative—faster responses were coded as more extreme attitudes). A highly positive or negative article followed about one of these hotels, supplemented by either a debunking or no fact-check. Participants then completed the evaluation task again. Dependent variable was persuasive impact: the difference between hotel evaluations before and after stimulus exposure (where negative values denote change incongruent to article valence).

Results and Discussion. A 2(debunking fact-check: yes/no) x 2(valence: positive/negative) ANOVA showed significant main effects of debunking, $F(1,380)=18.57$, $p<.001$, $\eta_p^2=.047$, and valence of the article, $F(1,380) = 64.34$, $p<.001$, $\eta_p^2=.147$. Debunking ($M=0.15$, $SD=1.36$) reduced the persuasive impact of the original message ($M=0.75$, $SD=1.41$), while negative articles ($M=0.86$, $SD=1.44$) were more persuasive than positive ($M=-0.18$, $SD=1.44$). In line with previous research (Nyhan & Reifler, 2010), debunking reversed persuasive impact for the positive message (from 0.06 to -0.32; $p_{diff}<.031$) and reduced, but did not reverse, impact for the negative message (from 1.37 to 0.58; $p_{diff}<.001$). No interaction between debunking and valence was found, $F(1,380)=2.25$, $p=.134$, $\eta_p^2=.006$.

---

[1] Additional conditions (both studies) and three measured variables omitted for space-related reasons. Pre-registered Study3 awaits completion.

Study 2

Method. Study 2 mimicked Study 1, but focused on medical clinics instead of hotels. 390 British participants (43.1% female, M_{age}=30.37, SD_{age}=10.49) were included.

Results and Discussion. Significant main effects were found for debunking, $F(1,249)$=45.85,p<.001, η_p^2=.156, and valence of the article, $F(1,249)$ = 66.24, p<.001, η_p^2=.210, on persuasive impact. Similar to Study 1, debunking reversed persuasive impact for the positive message (from 0.25 to -0.44; p_{diff}=.004), and reduced, but did not reverse, impact for the negative message (from 1.97 to 0.47; p_{diff}<.001). In terms of effect size debunking was, contrary to expectations, more effective for the negative message, $F(1,249)$=6.28, p<.05, η_p^2=.025.

General Discussion

Both studies suggest that debunking may be effective to fight the persuasive effects of misinformation, and they both show a similar pattern: debunking of negative information is effective, but some of its negative influence continues. In contrast, none of the positive influence continues when positive messages are debunked—rather, reading a debunked positive message effects peoples' evaluations negatively. Although these results are in line with prior claims that continued influence effects are asymmetric (Nyhan & Reifler, 2010), they were not in line with our own, preregistered, hypotheses. Thus, the results suggest we may have to update our understanding of corrections' persuasive effects.

Limitations. Possibly, both studies suffered from a ceiling effect (at T1, stimuli were evaluated above mid-point), causing the positive article to be relatively ineffective in elevating evaluations, in turn leaving little "room" for the debunking message to mitigate that effect. Future studies using the same design should use more neutral stimuli. Also, the negative messages used were mainly (moral) disgust evoking. Possibly, the negative messages invoked

emotional reactions focusing on risk avoidance, thereby inadvertently increasing negativity bias

(Baumeister et al., 2001).

References

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad Is Stronger Than

Good. *Review of General Psychology*, *5*(4), 323–370.

Ecker, U. K. H., Lewandowsky, S., Swire, B., & Chang, D. (2011). Correcting false information

in memory: Manipulating the strength of misinformation encoding and its retraction.

Psychonomic Bulletin and Review, *18*(3), 570–578.

Graves, L. (2013). *Deciding What's True: Fact-Checking Journalism and the New Ecology of

News*.

Lewandowsky, S., Ecker, U. K. H., Seifert, C. M., Schwarz, N., & Cook, J. (2012).

Misinformation and Its Correction: Continued Influence and Successful Debiasing.

Psychological Science in the Public Interest, Supplement, *13*(3), 106–131.

Nyhan, B., & Reifler, J. (2010). When corrections fail: The persistence of political

misperceptions. *Political Behavior*, *32*(2), 303–330.

Pan, L. Y., & Chiou, J. S. (2011). How Much Can You Trust Online Information? Cues for

Perceived Trustworthiness of Consumer-generated Online Information. *Journal of

Interactive Marketing*, *25*(2), 67–74.