

Title: Findings with only attentive participants appendix for “How effective are TikTok misinformation debunking videos?”

Authors: Puneet Bhargava (1), Katie MacDonald (2), Christie Newton (2), Hause Lin (1,3), Gordon Pennycook (1,2)

Date: March 29th, 2023

Note: The material contained herein is supplementary to the article named in the title and published in the Harvard Kennedy School (HKS) Misinformation Review.

Appendix A: Findings with only attentive participants

Thirty-three participants failed the last two attention checks. Similar to the main findings, below we report the findings after analyzing data of the remaining 1,136 participants.

Finding 1: Debunking (vs. misinformation-only) improves credibility of truthful TikTok videos.

We fitted a 2 (video veracity: true vs. false) x 2 (condition: debunking vs. misinformation-only) mixed ANOVA. Post-treatment true videos were found to be more credible than false ones, $F(1, 757) = 745.85$, $p < .001$, $d = 0.99$, but the credibility of post-treatment videos did not vary between conditions, $F(1, 757) = 1.27$, $p = .261$, $d = 0.04$. Crucially, as predicted, we found a significant interaction effect of veracity and condition on the credibility ratings of the subsequent TikTok videos, $F(1, 757) = 3.99$, $p = .046$, $d = 0.07$ (see Figure 4); that is, debunking improved overall accuracy for the subsequent video evaluation task. More specifically, debunking (vs. misinformation-only) increased subsequent credibility ratings for true videos, $b = 4.13$ (1.80), $t(1516) = 2.29$, $p = .022$, but had no effect on credibility ratings for false videos, $b = -1.48$ (1.86), $t(1516) = -0.79$, $p = .428$, leading to overall higher truth discernment in the debunking condition relative to the misinformation-only condition. This difference can be seen in Figure 5 (top).

Next, we fitted a 2 (video veracity: true vs. false) x 2 (condition: correction-only vs. misinformation-only) mixed ANOVA (Figure 4). True post-treatment videos were rated as more credible than false ones, $F(1, 758) = 634.00$, $p < .001$, $d = 0.91$, but credibility did not vary between conditions, $F(1, 758) = 0.40$, $p = .526$, $d = 0.02$, and the effect of veracity on credibility ratings of post-treatment videos, also, did not vary between conditions, $F(1, 758) = 0.17$, $p = .680$, $d = 0.01$. Thus, presenting only the correction video without the context of the original falsehood did not produce a significant increase in accuracy for the subsequent videos. Nonetheless, when we compared the debunking versus correction-only conditions with a 2 (video veracity: true vs. false) x 2 (condition: debunking vs. correction-only) mixed ANOVA, true (vs. false) post-treatment videos were found to be more credible, $F(1, 751) = 742.10$, $p < .001$, $d = 0.99$; but there were no differences in credibility or in the effect of veracity on credibility between conditions $F_s < 2.39$, $p_s > .122$, $d_s < 0.06$. Thus, although discernment was higher in the debunking condition relative to the misinformation condition (and this was not true of the correction condition), there was no difference between the debunking and correction conditions.

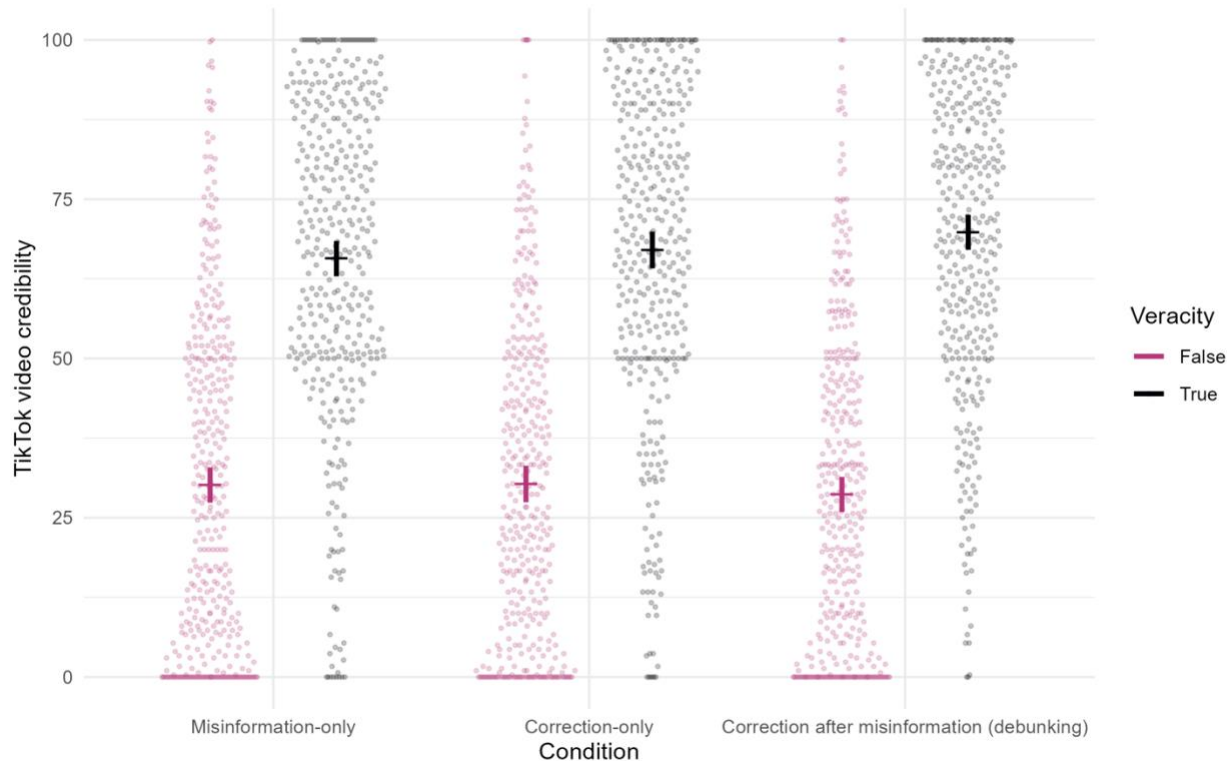


Figure 4. TikTok video credibility ratings for false and true videos and for the three experimental conditions. Credibility is computed by averaging reliability, unbiasedness, and accuracy ratings. Each dot is one participant's rating. Means and 95% CIs are shown.

Finding 2: Debunking (vs. misinformation-only) improves discernment, reduces false belief.

To further probe the effect of veracity on credibility ratings of post-treatment videos between conditions, we computed discernment scores for the video credibility ratings (credibility of true minus false videos (Figure 5 top). Replicating the mixed ANOVA results above, we find that, when collapsing across topics, discernment was higher in the debunking than misinformation-only condition, $b = 5.61 (2.84)$, $t(1133) = 1.97$, $p = .049$, $d = 0.12$; Figure 5 top left), but discernment was not significantly higher in the correction-only than misinformation-only condition, $b = 1.18 (2.84)$, $t(1133) = 0.42$, $p = .677$, $d = 0.02$ (Figure 5 top right). Although debunking increased discernment when collapsing across topics, the effects varied substantially across the six topics (Figure 5 top left).

In addition to rating the credibility of the TikTok videos, participants also indicated to what extent they agreed with or believed false statements associated with each TikTok video (Figure 5 bottom). Overall, participants in the debunking (vs. misinformation-only) condition believed the false statements less, $b = -0.37 (0.12)$, $t(1133) = -3.11$, $p = .002$, $d = -0.18$, and the effects were relatively consistent across topics (Figure 5 bottom left). Similarly, participants in the correction-only (vs. misinformation-only) condition also believed the false statements less, $b = -0.24 (0.12)$, $t(1133) = -2.06$, $p = .039$, $d = -0.12$, but the effect was smaller and varied more across topics (Figure 5 bottom right). Together, these results suggest that showing debunking videos on TikTok increased subsequent video discernment (but effects were small), whereas both showing debunking videos and correction decreased belief in false statements (effects were stronger for debunking than correction).

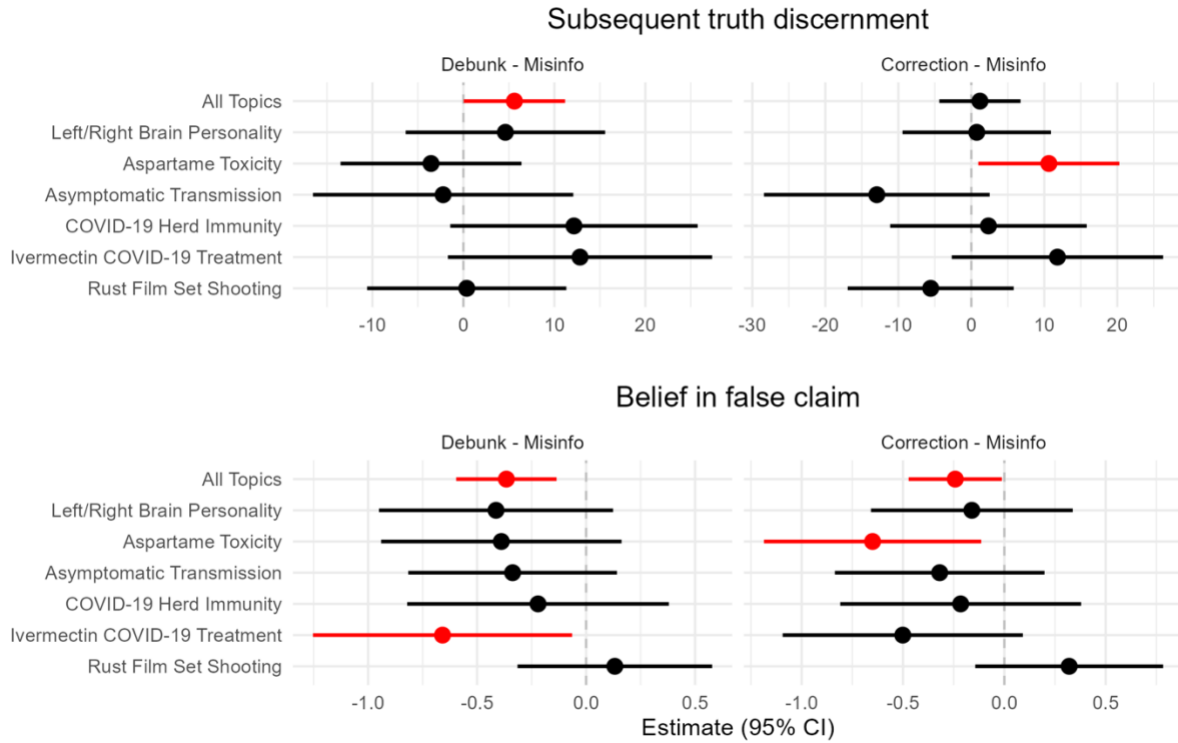


Figure 5. Effect of condition on TikTok video discernment (top) and belief in false statements associated with the TikTok videos (bottom). Discernment is the credibility ratings of true videos minus credibility ratings of false videos. Left panels contrast the debunking and misinformation-only conditions; right panels contrast the correction-only and misinformation-only conditions. The top row in each panel is the overall estimate for all six topics; the next six rows are the topic-specific estimates. Estimates highlighted in red are statistically significant ($p < .05$).