Prompting better sharing quality of COVID-related headlines

Irene Sophia Plank[1⋆]

1 Department of Psychiatry and Psychotherapy, LMU University Hospital, LMU Munich

⋆ irene.plank@med.uni-muenchen.de

Abstract

Asking people to evaluate the accuracy of a non-COVID-related headline decreases their likelihood to share COVID-related headlines, especially if they are false.

Received Date
Accepted Date
Published Date

Target article

A. A. Arechar, J. Allen, A. J. Berinsky, R. Cole, Z. Epstein, K. Garimella, A. Gully, J. G. Lu, R. M. Ross, M. N. Stagnaro, Y. Zhang, G. Pennycook *et al.*, *Understanding and combatting misinformation across 16 countries on six continents*, Nature Human Behaviour **7**(9), 1502 (2023), doi:10.1038/s41562-023-01641-6, Publisher: Nature Publishing Group

1 Goal

This robustness report investigates whether an alternative analysis supports the increase in sharing quality after encountering an accuracy prompt reported by [1].

2 Methods

Sharing behaviour is assessed using an online questionnaire with people rating how likely they would share a COVID-related article based on the headline (outcome *sharing likelihood* on a scale from (1) = 'moderately unlikely' to (6) = 'extremely likely'). Half of the headlines report false while the other half report true statements (predictor *Truth*). I define sharing quality as the difference between true and false headlines. Data was collected in 16 countries, using translated headlines. Some participants encounter an accuracy prompt before rating their sharing likelihood (predictor *Condition*). This prompt asks them to evaluate the accuracy of a single, non-COVID-related headline. The online questionnaire also contains attention checks and screeners. I excluded the data of 8,592 participants who failed at least one of these tests, whose total duration was longer than the 99% percentile (94.55 min) or who chose the exact same rating for all 20 headlines they encountered. For the included participants ($n = 8,587$), I

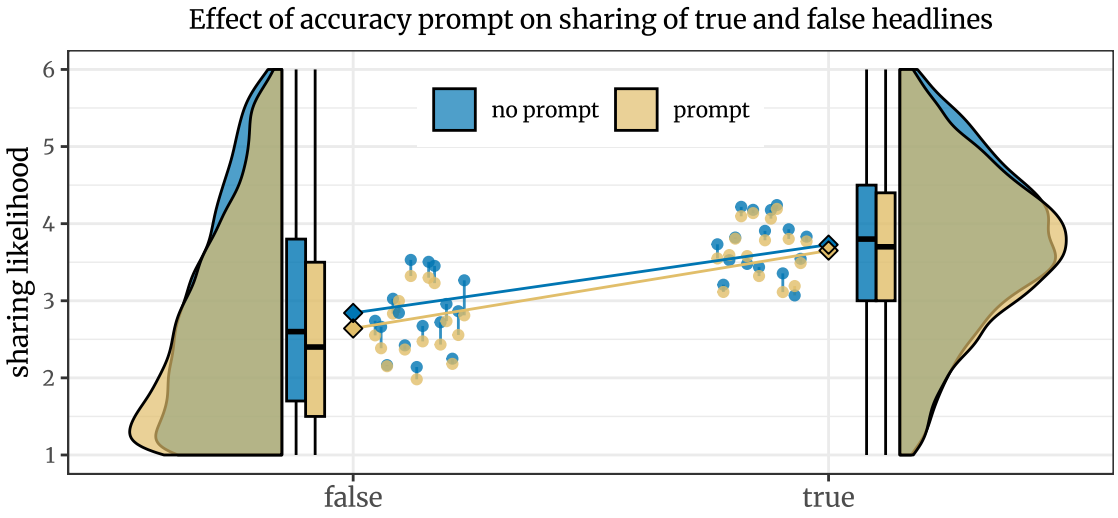

Figure 1: Density distributions and boxplots of participant averages. Dots show country averages with lines in the colour of the condition leading to a higher mean sharing likelihood. Diamonds show averages over all participants, regardless of country. Sharing likelihood was lower in the prompt condition for both true and false headlines, a pattern that was observed in almost all countries. However, this effect was slightly more pronounced for false headlines.

computed separate averages for their sharing likelihood of true and false headlines. These averages were modelled with a Gaussian Bayesian linear mixed model implemented in `brms` [2]. The model included two sum-coded population-level predictors, *Truth* (true or false headline) and *Condition* (prompt or no prompt), and their interaction as well as group-level intercepts for participant and country. Additionally, I included slopes for both population-level predictors and their interaction for each country. Posterior predictive checks revealed slight deviations in the overall shape but good capture of predictor means in the real data by predicted means based on the model. Standardised effect sizes δ were computed by dividing effects by the square root of the sum of all squared variances following [3].

3 Results

Overall, participants were more likely to share true than false headlines (*estimate* = 0.96 [0.85, 1.07], *posterior probability* = 100%; $\delta = 0.783$ [0.665, -0.897]). Participants who received an accuracy prompt before rating sharing likelihoods exhibited a credibly better sharing quality than participants who did not receive such a prompt, meaning that there was a greater difference between sharing likelihoods of true and false headlines (*estimate* = 0.12 [0.06, 0.18], *posterior probability* = 99.76%; $\delta = 0.098$ [0.036, 0.160]). Participants in the prompt condition rated their likelihood to share lower (*estimate* = -0.13 [-0.17, -0.08], *posterior probability* = 99.99%; $\delta = -0.103$ [-0.147, -0.059]), with this reduction applying to both true (*estimate* = -0.07 [-0.12, -0.01], *posterior probability* = 97.98%; $\delta = -0.054$ [-0.104, -0.002]) and false headlines (*estimate* = -0.19 [-0.24, -0.13], *posterior probability* = 100%; $\delta = -0.152$ [-0.208, -0.096]). However, the reduction was more pronounced for false headlines (see 1).

4 Conclusion

My results indicate that sharing quality was improved by a preceding accuracy prompt: while people rated their overall likelihood to share lower for both true and false headlines, this effect was more pronounced for false headlines. Thus, this re-analysis supports the original claim by [1] using a Bayesian approach.

Acknowledgments and Disclosures

Reproducibility We were able to computationally reproduce the original analysis and results.

Code and Data Availability Data and R code are available on OSF: https://osf.io/7wgv2/.

Funding The author received no financial support for the research, authorship, and/or publication of this article.

Conflicts of Interest There are no competing or conflicting interests.

References

[1] A. A. Arechar, J. Allen, A. J. Berinsky, R. Cole, Z. Epstein, K. Garimella, A. Gully, J. G. Lu, R. M. Ross, M. N. Stagnaro, Y. Zhang, G. Pennycook *et al.*, *Understanding and combatting misinformation across 16 countries on six continents*, Nature Human Behaviour **7**(9), 1502 (2023), doi:10.1038/s41562-023-01641-6, Publisher: Nature Publishing Group.

[2] P. C. Bürkner, *brms: An R package for Bayesian multilevel models using Stan*, Journal of Statistical Software **80**(1), 1 (2017), doi:10.18637/jss.v080.i01.

[3] L. V. Hedges, *Effect sizes in cluster-randomized designs*, Journal of Educational and Behavioral Statistics **32**(4), 341 (2007), doi:10.3102/1076998606298043, https://doi.org/10.3102/1076998606298043.