Accuracy prompts increase discrimination ability in news sharing but do not mitigate the effect of attitudinal congruence

Ilse L. Pit[1,2,3⋆]

1 Institute of Human Sciences, University of Oxford, United Kingdom
2 Calleva Research Centre for Evolution and Human Sciences, Magdalen College, United Kingdom
3 Center for Adaptive Rationality, Max Planck Institute for Human Development, Germany

⋆ ilse.pit@anthro.ox.ac.uk

Abstract

Accuracy prompts increased discrimination in headline sharing. Across multiple operationalisations, congruence had a large effect not mitigated by the prompt.

Received Date
Accepted Date
Published Date

Target article

Arechar, A. A., Allen, J., Berinsky, A. J., Cole, R., Epstein, Z., Garimella, K., ... & Rand, D. G. (2023). Understanding and combatting misinformation across 16 countries on six continents. Nature Human Behaviour, 7(9), 1502-1513. doi: 10.1038/s41562-023-01641-6

1 Goal

In this report I applied a signal detection theory framework to examine whether accuracy prompts improve the quality of news sharing, defined as the ability to discriminate true from false news by sharing the former and withholding the latter [1,2]. I also inspected the role of congruence, i.e. whether news aligns with a participant's prior attitudes, a factor known to affect both accuracy judgements [3] and sharing intentions [4].

2 Methods

I used participants' attitude towards COVID-19 vaccinations and categorized the headline content to calculate a score of congruence between the participant and headlines. I did this in 18 different ways, of which nine yielded samples suitable for the analysis. I modelled sharing (not likely to share = -0.5; likely to share = 0.5) using Bayesian multilevel probit regression using the function `brm` from the package `brms` [5] (version 2.22.0) in R (version 4.4.2) for each

of the operationalisations of congruence. Headline veracity (false = -0.5; true = 0.5), accuracy-prompt condition (control as reference), and headline congruence (incongruent as negative values; congruent as positive values) and two-way interactions to test whether prompt effects varied by veracity and congruence were included as fixed effects. I included varying intercepts for participant, headline, and country, and a varying slope of veracity for participants, and used weakly informative priors on the fixed effects. I synthesized the estimates across the different operationalisations of congruence using a random-effects meta-analysis implemented with the function `rma` from the package `metafor` [6] (version 4.8-0).

3 Results

As reflected in the meta-analysed estimates, across the 9 included operationalisations of congruence, participants showed no baseline tendency either to share or to withhold sharing headlines (b = -0.02, 95% CI [-0.12, 0.08]), and participants showed a positive baseline discrimination ability in their sharing decisions, i.e. sharing true headlines while withholding false ones (b = 0.73, 95% CI [0.65, 0.82]). Participants in the accuracy prompt condition were less likely to share headlines regardless of the veracity of the headline (b = -0.12, 95% CI [-0.14, -0.11]), and they had a higher discrimination ability as compared to participants in the control condition (b = 0.11, 95% CI [0.1, 0.13]). Participant shared more congruent headlines more as compared to baseline (b = 0.24, 95% CI [0.22, 0.25]), and showed a worse discrimination ability for more congruent headlines (b = -0.43, 95% CI [-0.48, -0.37]). There was no difference in how the accuracy prompt affected the overall sharing of more congruent headlines (b = -0.01, 95% CI [-0.02, 0.01]), nor the discrimination ability of more congruent headlines (b = -0.01, 95% CI [-0.03, 0.02]).

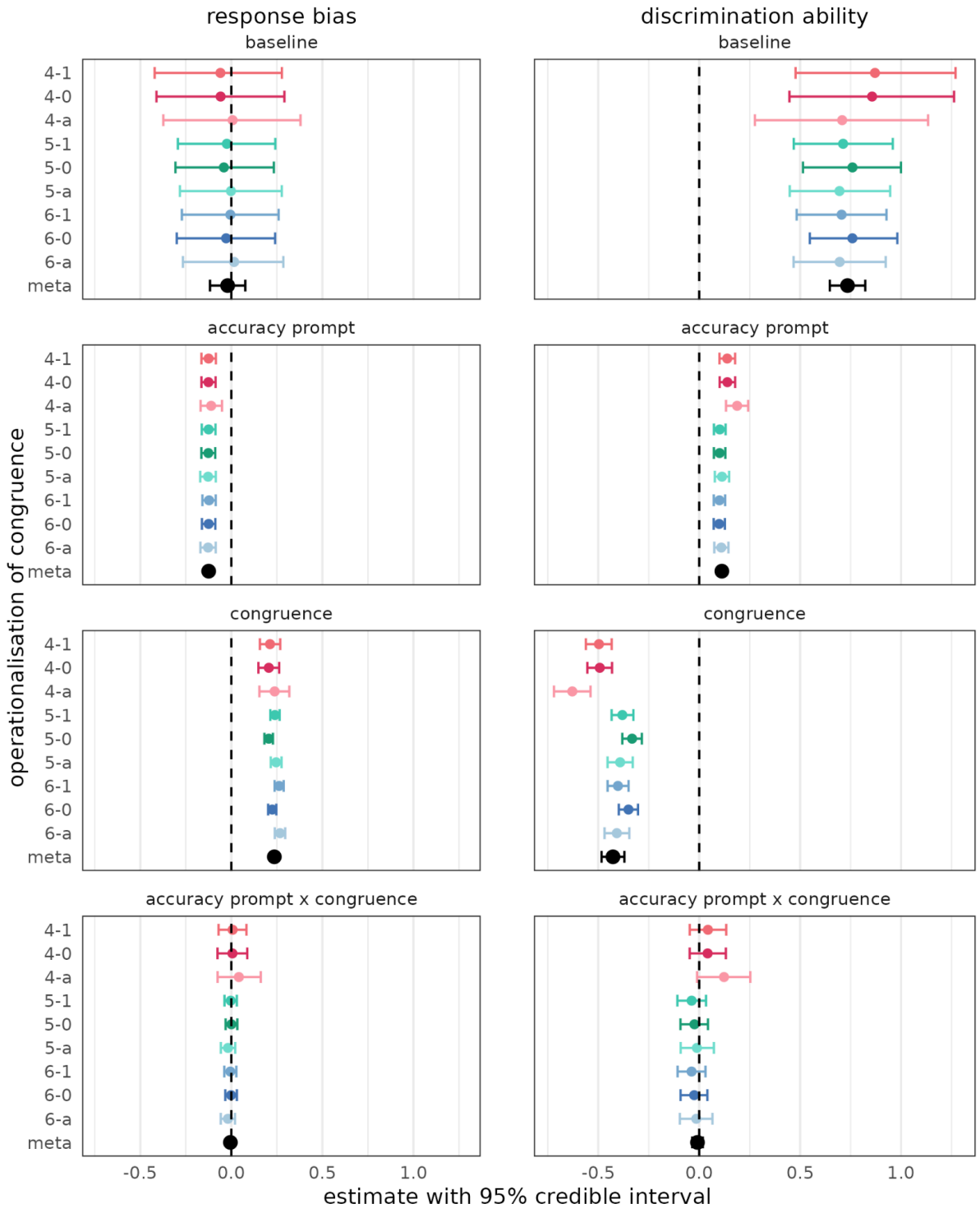

Figure 1: Estimates and 95% credible intervals for the baseline response bias towards true news and discrimination ability, and the additive effects of the accuracy prompt, headline congruence, and their interaction, for each of the different operationalisations of congruence. See supplementary information for details.

4 Conclusion

My findings corroborate Arechar and colleagues' [7] conclusion that an accuracy prompt improves the quality of headline sharing. The effect of the accuracy prompt on discrimination ability is relatively small compared to the baseline and the strong negative effect of headline congruence, which is not mitigated by the prompt. Despite this, the effect of the accuracy prompt on sharing quality is consistent within this report and with the authors' original findings.

Acknowledgments and Disclosures

Reproducibility I was unable to computationally reproduce the original analysis and results due to the use of licensed software unavailable to me.

Code and Data Availability Supplementary information can be found on the OSF at http://doi.org/10.17605/OSF.IO/4JHPT . This repository includes the data and codebook created to categorise the news headlines, code to create the different operationalisations of congruence and to reproduce the results and figure in this report.

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

References

[1] C. Batailler, S. M. Brannon, P. E. Teas, and B. Gawronski. *A signal detection approach to understanding the identification of fake news.* Perspect. Psychol. Sci. **17**, 1 (2022). doi:10.1177/1745691620986135.

[2] B. Gawronski, L. S. Nahon, and N. L. Ng. *A signal-detection framework for misinformation interventions.* Nat. Hum. Behav. **8** (2024). doi:10.1038/s41562-024-02021-4.

[3] M. Sultan, A. N. Tump, N. Ehmann, P. Lorenz-Spreen, R. Hertwig, A. Gollwitzer, and R. H. J. M. Kurvers. *Susceptibility to online misinformation: A systematic meta-analysis of demographic and psychological factors.* Proc. Natl. Acad. Sci. U.S.A. **121**, 47 (2024). doi:10.1073/pnas.2409329121.

[4] B. Gawronski, N. L. Ng, and D. M. Luke. T*ruth sensitivity and partisan bias in responses to misinformation.* J. Exp. Psychol. Gen. **152**, 8. (2023). doi:10.1037/xge0001381.

[5] P.-C. Bürkner. *brms: An R package for Bayesian multilevel models using Stan.* J. Stat. Softw. **80** (2017). doi:10.18637/jss.v080.i01.

[6] W. Viechtbauer. *Conducting meta-analyses in R with the metafor package.* J. Stat. Softw. **36** (2010). doi:10.18637/jss.v036.i03.

[7] A. A. Arechar, J. Allen, A. J. Berinsky, R. Cole, Z. Epstein, K. Garimella, A. Gully, et al. *Understanding and combatting misinformation across 16 countries on six continents.* Nat. Hum. Behav. **7** (2023). doi:10.1038/s41562-023-01641-6.