

Appendix B: Excluding participants who failed the attention check

There were two attention checks: one pre-treatment: “The colour test is simple, when asked for your favourite colour you must choose the word puce below. Based on the text you read above, what colour have you been asked to choose?” where participants had to choose between 5 colors, including “puce” and one post-treatment: “The colour test is simple, when asked for your favourite colour you must choose the word brown below. Based on the text you read above, what colour have you been asked to choose?” where participants had to choose between 5 colors, including “brown” (and not “puce”). In India, 292 participants failed at least one attention check, compared to 75 in Brazil and 90 in the United Kingdom.

In the pre-registration, we said that “We will re-estimate and compare all of our analyses by dropping individuals who fail our attention checks.” However, because it is not recommended to condition treatment effects on post-treatment variables, below we exclude only participants who failed the pre-treatment attention check.

We see that the effect sizes are similar when excluding participants who failed the attention check; however, the effects of corrections with a link in India are no longer significant, while the effect of correction without a link in India is significant. Beyond small differences in p -values (differences that are themselves not significant), these findings point in the direction that short user corrections may be slightly effective, but the effects are so small that very large sample sizes are needed to reliably detect them and that the added value of the link to fact checks is likely even smaller.

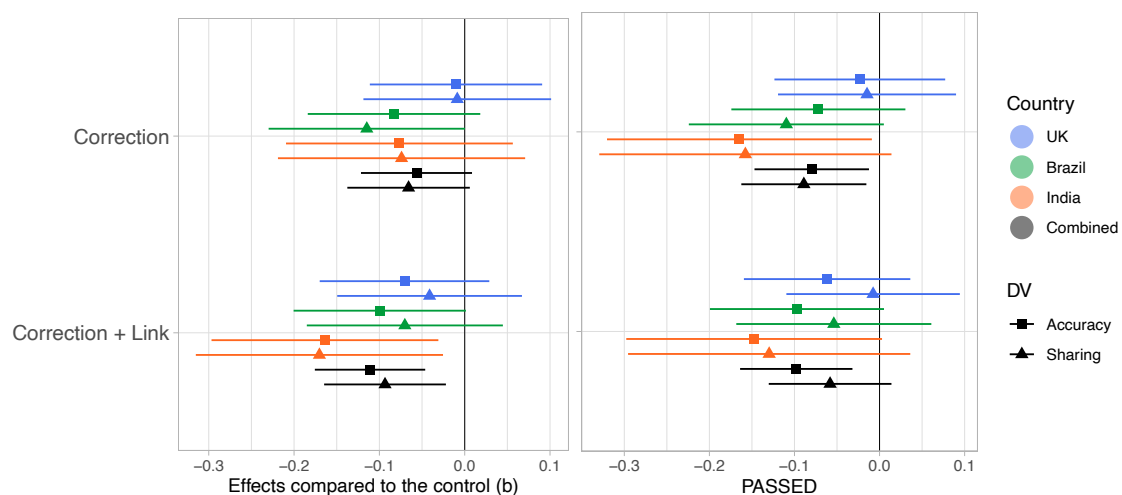


Figure B1. Effects of the corrections compared to the control on the 4-point scale (b). In the left panel, we do not exclude participants who failed the attention check, while on the right panel, we exclude participants who failed the attention check.