

More information on the sample restriction and data quality in both experiments

Main Experiment

The participants of the main experiment were recruited via CloudResearch's MTurk Toolkit which allows researchers to use multiple demographics and data quality filters such as workers' approval ratings, duplicate IP and Geocode block, workers' country location verification, CloudResearch-approved group of participants.

In our additional effort to improve data quality, as explained in the pre-analysis plan, we excluded participants who played the role of spectators if they (1) failed the Captcha; (2) did not pass the comprehension quiz (for which they only have one chance); (3) have already participated in the study; (4) spent too little time on reading the experimental instructions, making decisions, and completing the questionnaire (i.e., participants who completed the entire task in less than 2 minutes).

Overall, 71.4% (=1220/1708) passed all the exclusion criteria and are included for data analysis for spectators. Specifically, since Qualtrics appears to automatically screen out participants (potentially bots) who failed the Captcha, actually all recorded participants (N=1708) passed the Captcha. Each participant also has a unique Worker ID, which means that it is highly unlikely one worker has participated in the study more than once. Among them,

- 25.6% (=437/1708) did not pass the comprehension quiz
- 0.6% (=10/1708) did not complete the entire experiment
- 2.4% (=41/1708) completed the experiment in less than 2 minutes

Follow-up Experiment

The participants of the follow-up experiment were recruited via Connect which is CloudResearch's in-house platform that recruits participants independently from MTurk. Similar to their MTurk Toolbit, Connect also allows researchers to use multiple demographics and data quality filters such as workers' approval ratings, duplicate IP and Geocode block, workers' country location verification, CloudResearch-approved group of participants.

As a side note, there are multiple reasons why we switched to Connect for the follow-up experiment. One major reason is due to Amazon's sudden change in their billing policies which caused the freezing of one of the coauthors' MTurk account (which is required for using CloudResearch's MTurk Toolkit). We tried to contact Amazon to resolve this issue but in the end, Amazon was not responsive to our request in time before we had decided to switch to another platform. The second major reason is that there is a number of papers (see those cited in Douglas, Ewell, and Brauer, 2023) suggesting that CloudResearch and Prolific can offer better data quality than MTurk. Given that we are already quite familiar with CloudResearch and especially its capability to filter out participants with "bad" reputation as well as to target participants from

various demographic backgrounds, we eventually decided to use CloudResearch's in-house platform.

Douglas, Ewell, and Brauer (2023) also conduct an experiment comparing data quality between MTurk, Prolific, CloudResearch, Qualtrics and SONA. They show that Prolific and CloudResearch produce generally better data quality in terms of participants' likelihood to pass attention checks, provide meaningful answers, remember previous information, etc.

In our additional effort to improve data quality, as explained in the pre-analysis plan, we excluded participants who played the role of spectators if they (1) failed the Captcha; (2) did not pass the comprehension quiz (for which they only have one chance); (3) have already participated in the study; (4) spent too little time on reading the experimental instructions, making decisions, and completing the questionnaire (i.e., participants who completed the entire task in less than 3 minutes due to the longer survey).

Overall, 65.1% (=646/992) passed all the exclusion criteria and are included for data analysis for spectators. Note that the share excluded in the follow-up experiment is slightly higher than that in the main experiment. However, this is probably not surprising given the instructions of the follow-up experiment are also more complicated.

Specifically, since Qualtrics appears to automatically screen out participants (potentially bots) who failed the Captcha, actually all recorded participants (N=992) passed the Captcha. Each participant also has a unique Connect ID, which means that it is highly unlikely one worker has participated in the study more than once. Among them,

- 33.2% (=329/992) did not pass the comprehension quiz
- 1.0% (=10/992) did not complete the entire experiment
- 0.7% (=7/992) completed the experiment in less than 3 minutes

Reference

Douglas, B. D., Ewell, P. J., & Brauer, M. (2023). Data quality in online human-subjects research: Comparisons between MTurk, Prolific, CloudResearch, Qualtrics, and SONA. *PLOS ONE*, 18(3), e0279720.