Online Data Quality Study 2 (#59214)

Author(s)

Eyal Peer (Hebrew University of Jerusalem) - eyal.peer@mail.huji.ac.il

David Rotschild (Microsoft Research) - david@researchdmr.com

1) Have any data been collected for this study already?

No, no data have been collected for this study yet.

2) What's the main question being asked or hypothesis being tested in this study?

We compare four data quality aspects (attention, comprehension, reliability and honesty) between three sites used for online subject recruitment: Amazon MTurk, CloudResearch (formerly TurkPrime) and Prolific.

3) Describe the key dependent variable(s) specifying how they will be measured.

1. Attention is measured using two attention-check questions. The first asks participants to answer "six" and "three" to two items regardless of their actual preference (other responses are coded as failures); the second is an item within a scale worded "I currently don't pay attention to the questions I'm being asked in the survey" (response other than "strongly disagree" is coded as a failure)

2. Comprehension is examined through the oral summaries of instructions to two tasks. The first task asks participants to identify faces in a picture, but includes an instruction to only report zero; the second includes instructions for the "Matrix task" (Mazar et al., 2008). Two raters will code the oral responses independently and blind to the origin of the participant. We will consider an answer correct if both readers agree it is correct, and will apply a third reader to any answer with splits.

3. Reliability will be measured using Cronbach's alpha measure for the Need for Cognition scale.

4. Honesty will be measured using an online version of the Matrix task (Mazar et al., 2008) that will include two unsolvable matrices. Reporting solving any of these two problems will be coded as a dishonest response. Additionally, we will examine whether participants lie about their eligibility for a future study by asking them to indicate if they want to be invited to a study that samples participants of their own gender but whose age will be described as 5-10 years above the age participants reported in the beginning of the study.

In addition we will examine drop-out rates, duration for completing the survey, overall response time and speed between sites, differences in NFC, demographics, and patterns of usage of the site (main purpose, frequency of usage, number of submissions and approval ratings, usage of other sites), and we will also ask participants to report did they complete a study similar to this study in the last months.

4) How many and which conditions will participants be assigned to?

No conditions, the comparison is between the three sites mentioned above.

5) Specify exactly which analyses you will conduct to examine the main question/hypothesis.

We will use chi-square tests to examine differences in the rates of attention, comprehension and dishonesty. We will also compute average scores for (across the items per aspect) and compare them between sites using ANOVA and regression analyses. We will test for differences between reliability coefficients using Hakistan & Whalen (1976) method. Lastly, we will compute an overall composite score of data quality (based on the average scores of attention, comprehension and dishonesty) and compare them between sites using ANOVA and regression.

Depending on how many participants will report they have done a similar study in the past, we may compare the above data quality aspects between naive and non-naive participants.

6) Describe exactly how outliers will be defined and handled, and your precise rule(s) for excluding observations.

We will exclude participants who did not complete all of the study (after using their responses to calculate drop-out rates).

We will also restrict the samping on all sites to participants who did not already complete a previous similar study we conducted using a "black list" of their worker/participant IDs.

7) How many observations will be collected or what will determine sample size? No need to justify decision, but be precise about exactly how the number will be determined.

We will post the study asking for 500 participants from each site. However, we will retain any additional responses, if such are collected. Participants will be only U.S. residents age 18 or above.

8) Anything else you would like to pre-register? (e.g., secondary analyses, variables collected for exploratory purposes, unusual analyses planned?)

We will apply data quality filters on all three sites in the following scheme:

1) MTurk - only workers who have completed 100 submissions or more and 95% of those submissions were approved.

2) CloudResearch - only workers who have completed 100 submissions or more and 95% of those submissions were approved + the default setting on CloudResearch of "block low data quality workers"

3) Prolific - only participants who have completed 100 submissions or more and 95% of those submissions were approved.