

Transparency in Dependency on Author Identities (at ITCS 2023) (#105722)

Created: 08/29/2022 10:52 AM (PT)

Public: 12/13/2022 02:34 PM (PT)

Author(s)

Nihar Shah (CMU) - nihars@cs.cmu.edu

Yael Kalai (MSR) - yael@microsoft.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?

How do the scores and confidences of reviewers change after seeing author identities?

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

Scores given by reviewers in their reviews and the confidence self-reported by reviewers.

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

A single condition -- there is no randomization. All participants (reviewers) will first be shown anonymized papers, and once they submit their review, they will be able to see author identities and allowed to change their review.

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

We will conduct two primary analyses.

(1) First, we will show the distribution of change in self-reported confidence and change in scores given. We will specifically point out the mean and median change.

(2) For each review included in the analysis, we will compute the rank of the affiliation of the associated paper's authors as follows. For each author, we will compute their affiliation's rank using csrankings.org under the theory filter. In case an affiliation is not covered by csrankings.org, we will elicit the expertise of the program chair on the ranking of that affiliation in the list of all affiliations. If there are multiple affiliations associated with any author, we will take the best of these ranks. We will then compute a rank associated to each paper as the best rank among all authors. This will give us a partial ranking of all papers.

We will then compute the correlation between the change in scores and self-reported confidence with the ranking of papers' affiliations. The correlation will be computed in terms of Kendall's tau b correlation coefficient.

As a secondary analysis, we will also stratify according to responses to the question of whether the reviewer knows the identity of the authors. Here, we will consider two responses separately -- no knowledge of author identities and having a guess about author identities.

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

We anticipate the following types of outliers.

1. We have asked participants to indicate if they know authors' identities. If they say yes, then we will exclude this data.
2. Some participants may submit a blank or placeholder review in order to see author identities and then write their review. We will also exclude this data.

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.

This analysis is conducted in the ITCS 2023 conference. The paper submission deadline is September 8, 2022 after which the papers will be reviewed. The number of samples will equal the number of reviews, minus the aforementioned exclusions.

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

An exploratory analysis of the change in the text of the reviews before and after reviewers see author identities.

A survey of participants after the reviewing activity about their perceptions on anonymizing author identities.