

## AI Language Discernment - Heuristics validation study (#90596)

Created: 03/11/2022 02:14 AM (PT)

Public: 02/03/2023 04:03 AM (PT)

### Author(s)

Maurice Jakesch (Cornell University) - mpj32@cornell.edu

### 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?

The current experiments validate the findings of a prior correlative study where we examined which language features make people believe that personal self-descriptions were generated by AI. We use features that were predictive of perceived humanity in the previous study to build a pipeline that algorithmically generates self-descriptions optimized for perceived humanity. We then test whether self-descriptions with language features that were correlated with perceived humanity in the previous study indeed produce an impression of humanity in an independent experiment.

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

Perceived AI generation on a 5-point Likert-scale ranging from "The profile was written / generated by (1) Definitely AI ... (5) Definitely human".

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

Each participant evaluates 16 self-descriptions five of which are real, human profiles. Another five are regular AI-generated profiles and the final six are AI-generated profiles optimized for humanity leading to a 3-factor within-subject design.

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

Linear regression predicting what percentage of participants will believe that a self-description was generated based on its type (real, generated, generated and optimized for humanity).

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

N/A

### 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 recruit 100 participants in each context (dating and professional profiles), so 200 overall.

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

Nothing else to pre-register.