Robin Murphy (robin.murphy@psy.ox.ac.uk)  
Riddhi Pitliya (riddhi.jain@psy.ox.ac.uk)  
Daniel Lametti (daniel.lametti@acadiau.ca)

# **1) Data collection.** Have any data been collected for this study already?

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

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

The main hypothesis is that people affiliate more with Large Language Model (LLM) agents that resemble their own personality. We predict that participants will have higher affiliation scores with an LLM prompted to produce language that mirrors their personality (mirror condition) compared to an LLM prompted with the inverted scores of their personality (inverse condition).

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

The key dependent variable is the affiliation score, which is an average metric based on six Likert scale questions: “I felt that we are similar”, “I enjoyed our conversation”, “I would chat with them again”, “I felt that they were different from me” (inverted), “I felt distant from them” (inverted), “I felt that they understood me”. As described in Castiello et al., 2025 – PsyArXiv.

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

Participants will be exposed to two text-based conversations with two different LLMs. The first LLM will be based on the participant's Big-5 personality traits, and the second (inverted) will be based on the inverse of these traits. Initially, participants will complete the 44 item Big-5 personality assessment. Then, they will be randomly assigned to either the mirror condition followed by the inverse condition or vice versa. This crossover design ensures that each participant experiences both conditions. Each condition is entirely dependent to the questionnaire score for each participant.

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

We will conduct a paired t-test between the mirror and inverse conditions, setting the alpha level at .05.

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

Participants who complete fewer than 8 conversational turns in a chat or fail the anti-AI attentional checks will be excluded from the study. Additionally, we will exclude participants with identical personality scores for the mirror and inverse conditions, i.e., complete neutrality, making both conditions the same by scoring an average of 3 (“neither agree nor disagree”) on all Big-5 dimensions.

**7) Sample Size** How many observations will be collected or what will determine sample size?

We calculate the sample size with function pwr.t.test from the R package pwr. Based on a previous pilot, we need 90 participants to detect an effect size of .3 with a power of .8 and an alpha of .05.

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

We will run a liner model predicting the affiliation score using condition as binary regressor, the distance between LLMs’ personalities, and the interaction. A significant interaction would suggest a dose-response effect between condition and affiliation. The distance between LLMs’ personalities would moderate the main effect between conditions. The distance between LLM’s personalities is the dot product between the 5-dimensional vectors.

**9) Name** Give a title for this AsPredicted pre-registration

Human-LLM affiliation based on the presentation of shared BFI traits

homophily between artificial personality and humans

# **10) Type of study.**

Experiment

# **11) Data source**

Prolific