

The effect of AI conversation on polarization (#234546)

Author(s)

This pre-registration is currently anonymous to enable blind peer-review.
It has 2 authors.

Pre-registered on: 2025/06/20 - 05:23 AM (PT)

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?

Does having a conversation with an empathetic AI that presents opposing viewpoints, reduce polarization among individuals with strong, polarized beliefs?

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

Polarization level will be measured through Pre- and Post-survey questions asking participants to rate agreement on divisive (political) issues.

Pre-survey: Participants will rate their belief with polarizing statements (e.g. on a Likert scale)

Post-survey: After AI conversation, participant will rate their belief in a specific polarization statement.

Dependant variable: (Political) polarization, which will be measured by assessing changes in participants' political beliefs before and after the empathetic human-AI conversation.

Pre/Post-survey questions (all rated on a scale from 0%(strongly disagree) to 100%(strongly agree)):

"Human activity is the main cause of global warming."

"The U.S. benefits from immigration and should avoid extreme border restrictions."

"Gun laws should be stricter in the U.S."

"Replacing the federal income tax with tariffs is a fair way to fund the government."

"The 2024 U.S. election results were legitimate and trustworthy."

"AI will make human jobs obsolete and that's a dangerous path."

"Health mandates like vaccines and lockdowns are sometimes necessary and justified."

"The U.S. foreign policy contributes to peace in the Middle East."

"'Woke' culture is essential for achieving true social justice."

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

All participants will complete the same AI conversation task. There is only one condition.

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

We predict that participants will show lower polarization scores after the empathetic AI conversation. To test this, we will conduct a paired t-tests comparing polarization scores from pre- and post-survey.

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

Participants will be excluded from the analysis if they fail the pre-survey on polarization. Those who do not express sufficiently strong or polarized views based on the pre-survey.

Additionally, participants will respond to an open-ended question that serves as an attention check. This will not be used for exclusion but may be used for later data analysis.

We aim to recruit people from the US, fluent in English, ≥ 18 years, with high Prolific approval rating.

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 aim for at least 100 participants.

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

10-day follow-up:

Participants will be invited to complete a 10-day follow-up survey to assess any long-term changes in their polarization levels after the initial interaction with the empathetic AI. The follow-up survey will include the same polarization measures as the pre- and post-surveys to evaluate whether the reduction in polarization persists over time.