### **Study Information**

1. **Title:** Perceptions of artificial intelligence systems
2. **Description**

The primary purpose of this study is to look at how people think about AI in general, and how this maps on to perceptions of other specific AI systems. We plan to determine which kinds of AI systems are perceived as more “prototypically AI” and whether this prototypicality varies based on different measures.

A secondary purpose of the study is to do an exploratory “data dive” into the correlations between trust in AI systems and different variables discussed in the literature on AI (e.g., perceived ethicality, competence, autonomy, explainability) but which have not been studied together with a large number of AI types. These exploratory analyses and exploration of the different AI types will provide testable hypotheses for subsequent research with a shorter survey length.

### **Design Plan**

1. **Study type**

Online survey

1. **Blinding**

There is no blinding in this study.

1. **Study design**

This is descriptive and correlational study. The survey includes questions about (1) artificial intelligence in general, and (2) several specific types of AI.

1. **Randomization**

Participants will be randomly presented with a subset of AI types and predictor measures (see “Manipulated and measured variables” below).

### **Sampling Plan**

1. **Existing data**

Registration prior to creation of data: As of the date of submission of this research plan for preregistration, the data have not yet been collected, created, or realized.

1. **Data collection procedures**

Participants will be recruited through Prolific Academic. Data will be collected online via Qualtrics survey forms.

1. **Sample size**

We will recruit 800 participants.

1. **Sample size rationale**

We extrapolated our target sample size from a power analysis. The power analysis suggested that a sample size of 85 participants would be necessary to detect a medium correlation (*r* = 0.3) with 80% power. We aimed to achieve this sample size for every AI type and predictor measure combination. Given that participants see a random subset of AI types (5 out of 20) and a random subset of predictor measures (5 out of 11), we extrapolated the overall sample size to 85 x (20/5) x (11/5) = 748 participants. Because some participants may fail attention checks or not fully complete the survey, we will round this number up to N = 800 to maintain the required power.

1. **Manipulated & measured variables**

Manipulated

There are no experimental manipulations in this study. However, we will present participants with a random subset of five AI types from the following list:

* Skin cancer diagnosis app
* AI therapist
* Medical triage AI
* Facial recognition AI
* Predictive policing algorithm
* Predictive sentencing algorithm
* ChatGPT
* DeepSeek
* Audio transcription AI
* Instagram filter
* DALL-E
* Military cybersecurity AI
* Autonomous killer drone
* Robot soldier
* Robot vacuum
* Apple's Siri
* Google Maps AI
* Air traffic control AI
* Self-driving car
* AI superintelligence

Outcome measures

All participants will see the following primary outcome measure, on a 1-7 Likert scale, for both general AI and specific AI types:

* **Trust**: “How much do you think [X] can be trusted?”

Predictor measures

For general AI and specific AI types, participants will see a random subset of five predictor measures (on 1-7 Likert scales) from the following list:

* **Reliable:** “How reliable do you think [X] is?”
* **Competent:** “How competent do you think [X] is?”
* **Genuine:** “How genuine do you think [X] is?”
* **Ethical:** “How ethical do you think [X] is?”
* **Autonomy:** “How much freedom do you think [X] has in choosing how to behave?”
* **Potential good:** “How much do you think [X] has a potential for good?”
* **Potential harm:** “How much do you think [X] has a potential for harm?”
* **Interpretable:** “When [X] does something, how much do you think you could understand why it does that?”
* **Explainable:** “How much do you think someone could explain the inner workings of [X]?”
* **Human-like:** “How human-like do you think [X] is?”
* **Predictability:** “How much do you think you could predict the behaviour of [X] in a new context?”

Additional measures

For each specific AI type, we will also ask the following Yes/No questions:

* **Heard**: “Have you heard of [X] before?”
* **Used**: “Have you ever used [X]?” (not asked for AI superintelligence)

Individual-level and demographic measures

At the end of the survey, we will ask all participants:

* **AI familiarity**: “Overall, how familiar are you with AI tools?” 1-7
* **AI frequency of use**: “Overall, how frequently do you use AI tools?” 1-5
* **Generalised trust**: “Generally speaking, how much do you think that people can be trusted?” 1-7

We will then ask all participants the following demographic questions: age, gender, education, subjective SES, political ideology, and religiosity.

### **Analysis Plan**

1. **Analysis plan**

As a first step, we plan to use principal components analysis to determine which kinds of AI systems are perceived as most “prototypically” AI. For general AI and for each specific AI type, we will take the average scores for all outcome and predictor variables (e.g., trust, reliability, competence, etc.). We will then use principal components analysis to reduce the dimensionality of the dataset to fewer principal components. The exact number of components will be decided using a scree plot. We will then plot where the specific AI types sit in the multi-dimensional space and compare their distances to the general AI to assess prototypicality.

Beyond our initial analysis, we will also run further exploratory analyses looking at the correlates of trust in general AI. We will assess which predictors have the strongest correlations with trust using a Pearson’s correlation matrix.

1. **Data exclusion**

Participants who close their browser window without completing the experiment will be excluded.

Participants with low Captcha scores (<0.5) will be excluded.

Participants will also be excluded if they fail the following attention check at the beginning of the study:

*In studies like ours, there are sometimes a few people who do not carefully read the questions they are asked and just "quickly click through the survey." These random answers are problematic because they compromise the results of the studies. It is very important that you pay attention and read each question. In order to show that you read our questions carefully (and regardless of your own opinion), please answer "TikTok" in the question on the next page.*

[page break]

*When an important event is happening or is about to happen, many people try to get informed about the development of the situation. In such situations, where do you get your information from?*

Options: [TV, Twitter, Radio, Facebook, Youtube, Newspapers, Reddit, TikTok, Other]

1. **Missing data**

The survey follows an incomplete design: participants are randomised to only see a subset of predictor measures and a subset of AI types. As a result, there will be “missing data” for predictor measures and AI types that are not shown. However, this is not an issue for our planned analyses – we will not need to exclude data points on the basis of missing data.