Pre-Registration (and other best practices) Tutorial

HRI Lab Meeting 5/22

Transparent research

- There is a general goal/standard in research to achieve a significant ($p < .05$) result in order to get published or recognized

- But this emphasis on the significant result can lead to not great practices: "p-hacking"
  - Don't report insignificant results
  - Run multiple tests to find a significant value, even if this was not the initial plan for the study
  - In worst cases, researchers manipulate or falsify their data to get a significant result

Transparent research

- This can lead to misinterpreted or fallible results
  - Which can then lead to a general mistrust in scientific research
  - Also can further a cycle of bad research (e.g., you attempt to replicate a finding but cannot because the initial finding was inaccurate).

Transparent research

- Psychology research has taken a number of steps to promote transparent research
  - Pre-registering their study plan, hypotheses, and tests
  - Adding p-value corrections based on the number of tests ran
  - Including the raw data and analyses in a repository for other researchers to replicate

What is a Pre Registration?

- Before the study is ran (or was ran but before analyses have been conducted), the authors submit a document that details the following…
  - Hypotheses of the study
  - Design of the study
  - Sample size and justification
  - Participant exclusions
  - Analysis plan
  - Other disclosures

- There are many ways to submit a preregistration, but my personal favorite is using AsPredicted

Preregistration tips in general

- It should be understandable to any person

- Streamlined as much as possible

- Strike the right balance between general and specific with your analysis plan

- Mention all possible tests you could see yourself doing

- Justify your sample size
  - Either through prior work or power analysis

Beliefs about a robot's moral action (#152816)

Author(s)

This pre-registration is currently anonymous to enable blind peer-review.

It has 4 authors.

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?

This study investigates how U.S. adults respond to a robot's moral transgression compared to another person and whether this depends on the description of the robot's capabilities. We suspect that adults will generally be less likely to view the robot as a moral agent (e.g., less punishable, less intentional) than a person. However, when a robot is described as having mental or social capabilities, adults will view the robot as more like a moral agent than if the robot is described as only having physical capabilities.

main question is understandable to someone unfamiliar with this work and hypotheses are general, but specific to the study

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

Adults will be asked about their beliefs about the agentic capabilities of the robot/person on a scale from 0-3 (0=definitely no, 1 = somewhat no, 2 = somewhat yes, 3 = definitely yes): hunger, pain, fear, tiredness, guilt, embarrassment, love, anger, thought, sensation, memory, and reasoning.

After watching the video of the robot/person push somebody, adults will be asked about the moral agency of the robot/person on a scale from 0-3 (0=definitely no, 1 = somewhat no, 2 = somewhat yes, 3 = definitely yes): moral knowledge, emotional knowledge, care, moral desire, moral intent, moral punishment, moral personality.

Adults will also be asked about the robot/person's free will in pushing (binary). Adults will be asked to explain their answer. Open ended responses to this question will be categorized by two experiment-blind coders.

Adults will also be asked open-ended questions about how the robot/person felt about pushing the victim and how the victim felt about being pushed. Open ended responses will be categorized by two experiment-blind coders.

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

Adults will be randomly assigned to one of 16 between-subjects conditions in a 2 (Transgressor) by 4 (Information) by 2 (Counterbalance) design. The Transgressor will either be a robot or human. The Information given will either be about the transgressor's mental capabilities, physical experiences, socio-emotional capacities, or no information at all. The Counterbalance will either have adults complete the agentic capability questionnaire at the beginning of the study (after information given, before transgression video) or at the end of the study (after moral questionnaire).

Be more specific here about your study
design

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

To investigate whether adults judged the moral transgression differently between conditions, we will run separate GLMs with the response to moral agency questions as the DV and with Transgressor condition, Information condition, and Counterbalance condition as the IVs. We will also look at how responses differ for each question by running a repeated measures GLM with the response as the DV, question type, Transgressor condition, Information condition, and Counterbalance condition as the IVS, and participant ID as the random variable. We will first include interactions between conditions but remove them if they are found to be insignificant.

We will also run a logistic model for the free will question, with response as the DV and Transgressor condition, Information condition, and Counterbalance condition as the IVs.

We will also compare adults' agency judgments by running separate GLMs with the response to the agency questions as the DV and Transgressor condition, Information condition, and Counterbalance condition as the IVs. We will also look at how responses differ for each question by running a repeated measures GLM with the response as the DV, question type, Transgressor condition, Information condition, and Counterbalance condition as the IVS, and participant ID as the random variable.

For all of these models, we will first include interactions between conditions but remove them if they are found to be insignificant.

Here you want to be specific on the type of test and the variables you are using, but be general with any variables that you are less certain on (e.g., interactions). Include multiple tests (e.g., run both separate GLMs and repeated measures GLM)

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

We will exclude participants if they failed one or multiple attention checks placed throughout the study.

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 stop data collection once we have collected 720 participants (45 for each condition) that have completed the study without complications (see above response for exclusion criteria). Conducting a power analysis determined 45 per condition would be enough to get a small effect size (0.1). We may collect additional participants if our original sample is not representative of different demographics (e.g., race, gender).

Be specific on how you will exclude participants, your number of participants, and your justification for that number. But you can be general if you anticipate any changes

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

If we do not find any differences between conditions in our dependent variables, we will run equivalence tests.

Available at https://aspredicted.org/vk7f-k86d.pdf

Version of AsPredicted Questions: 2.00

ASPREDICTED

Wharton
CREDIBILITY LAB
UNIVERSITY of PENNSYLVANIA

We will explore any effects of demographic variables (e.g., age, gender, race) on the DVs.

We will run an exploratory factor analysis on the agency questions and run follow-up comparisons on the found factors (e.g., compare condition differences).

We will look at the relationship between the agency and moral agency responses.

Report any exploratory analyses you may anticipate doing. If you have also already begun data collection, you can explain here why you still consider this to be a preregistration (e.g., ran 5 participants and looked at means, but have yet to do any significance tests)

What if you want to run new analyses?

- That's okay! The preregistration is to improve transparency with you and the reader, not to restrict your research

- If you have not finished data collection, but decided on an entirely new analysis plan, you can submit a new preregistration (but give reasons why you are doing a new one)

- You still need to run and report the registered analyses (but you can report these in a supplementary materials)

- Provide justification for why you are conducting a different analysis
  - E.g., "Upon further review, the authors determined that a logistic model was a more accurate test for the variables than a linear model."

- Specify that the analyses were not registered
  - E.g., "In an exploratory analyses, we also looked at…"

How to report the preregistration in your paper

In methods

- Intro sentence saying the study design, hypotheses and participant size were preregistered (include the link)
- Mention any variations from the preregistration if you have any (e.g., collected an additional x number of participants)

In results

- When outlining the results, include a sentence that the analyses were preregistered
- If you conduct any results that were not registered, you must describe the findings as unregistered or exploratory

Where to put your preregistration

- Include it in your study's repository (e.g., Open Science Framework)

- In the repository, it also is best practice to include
  - Dataset (anonymized)
  - Analysis code
  - Any other programming code you may have used
  - Any stimuli used

Other things to keep in mind

- You don't have to preregister every study
  - If you have an exploratory, open-ended study, it is normal to not register
  - If you have multiple studies in one project, it helps having the first one be exploratory, and then you can preregister the follow-up studies
- Ideally, a preregistration shouldn't be too much extra work, as you have likely already thought of these questions before you even run the study
- When determining your sample size, you can base it off of prior work or a power analysis (I can go over this if we have time)