

Experiment LLM: Documentation

Tehilla Ostrovsky

2023-07-23

Demo-Experiment

The experiment will be about a simple perceptual task, in which participants have to judge a stimuli and will face the trade-off between speed and accuracy.

The experiment will include:

1. Multiple trials
2. Multiple conditions of difficulty (ideally 2-3)
3. Two options for response: Button press and Keyword
4. Feedback about the response. For example, a “correct”, incorrect

The experiment-text will be converted to LLM for analysis:

prior to LLM:

1. data cleaning
2. text validation:
 - 2.1 **Text Validation based on stimulus** - Ask 1/3 of the participants to describe the colour of the stimulus.
 - 2.2 **Text Validation based on condition (difficulty)** - Ask 1/3 of the participants to describe the level of the difficulty of the trial.
 - 2.3 **BONUS** - Ask 1/3 the participants to describe their confidence on each trial.
 - 2.4 **BONUS addition** - find a way to find words in the text that are inline with confidence such as “im sure”, “it must be...”

For 2.1-2.4 - Here is a blog that does that with LLM on Keyword Extraction with BERT: <https://jaketae.github.io/study/keyword-extraction/>

NOTES:

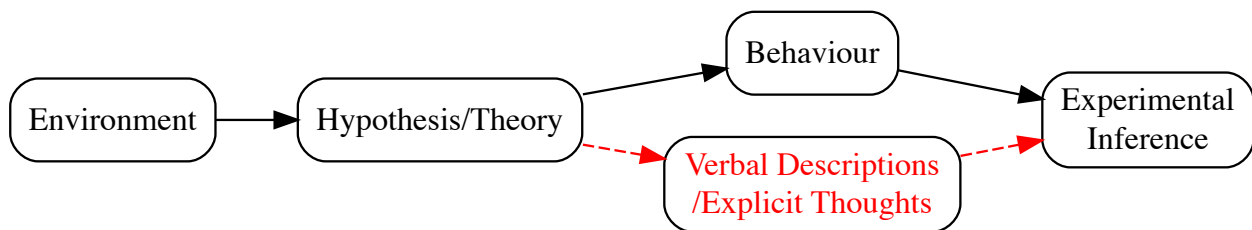
For the paper:

- add quality score of the recording - this is a problem unless we could establish the known WER (word error rate) see wiki on WER
- add 1 GPU instruction
- add potential fine-tuning of the LLM model ?

Experiment + Data collection

Outline for the Psych-Review Paper

1. The Interaction Between Environment (E), Theories about the Environment (T), and Behaviour (B)



1. Model's complexity and its interaction with verbal descriptions.

1.1. If the model/theory is **hard** to vary → the model/theory provides a good explanation to behavior = any additional explanations or variations will a poorer account for the data. Verbal descriptions sit tightly together with the explanation/theory proposed by the model.

1.2. “That is a good explanation – hard to vary, because all its details play a functional role.(The beginning of infinity)”

1.3. If the model/theory is **easy** to vary → there are more possibilities for the model to be wrong, that is, a model that is more flexible, will allow for more potential untested/non-explicit assumptions to slip in.

→ These assumptions can be revealed or refuted based on verbal descriptions.

1.4 - Experimenters often only lay the possibilities of untested, alternative post-hoc explanations in their discussion sections, never really engaging with them. ! Find examples for this idea

“Good explanations in science are hard-to-vary assertions about reality. They are hard-to-vary because they provide specific details that fit together so tightly that changing them ruins the explanation. This criterion helps to eliminate bad explanations that keep adding justifications in light of refutations and counterevidence to avoid falsification. An explanation that is hard-to-vary but does not survive a critical test can be considered falsified.”

“Good explanations help us achieve better map-territory convergence. They allow us to construct more accurate models of the territory.”

“Poor explanations purport to explain anything and everything. Such explanations explain nothing.”

2. Model's untested assumptions and verbal descriptions.

2.1. Verbal descriptions as potential tool for holding models accountable for their untested assumptions (ideas/hypotheses tested by participants during the experiment). 2.2 Verbal descriptions can help experimenters to specify if their tested models are expected to make predictions if people have some access to the processes they describe. For instance, does the DDMs assumptions of drift rate, starting bias etc are phenomena that people can describe.

Hence,

3. Verbal descriptions as a complementary data to validate models

Verbal descriptions are richer data source relative to

1. EEG
2. Eye tracking
3. Mouse Tracking
4. fMRI