The Mixed Subjects Design:
Treating Large Language Models as (Potentially) Informative Observations

Abstract: Large Language Models (LLMs) promise to transform the social sciences through cost-effective predictions of human behavior. However, despite growing evidence that LLMs can misrepresent such behavior, current approaches to studying causal effects with LLMs require researchers to assume that predicted and observed behavior are *interchangeable*. Instead, we argue that human subjects should serve as a gold standard to correct misrepresentations within a *mixed subjects design*. This paradigm offers valid and more precise estimates of causal effects at a lower cost than experiments relying solely on human subjects. We demonstrate—and extend—prediction-powered inference, a statistical method that instantiates the mixed subjects design. Our innovation is a power analysis for optimally choosing between *informative but costly* human subjects and *less informative but cheap* predictions of human behavior. Mixed subjects designs could enhance scientific productivity and reduce inequality in access to costly evidence on research questions by offering valid, precise, and cost-effective inferences on causal effects and other parameters.

Authors

David Broska is a Ph.D. candidate in Sociology at Stanford University.

- dbroska@stanford.edu (Corresponding Author)

- `https://orcid.org/0000-0002-5176-1163`

- Department of Sociology
  Stanford University
  450 Jane Stanford Way
  Stanford, CA 94305âĂŞ2004 USA

Michael Howes is a Ph.D. candidate in Statistics at Stanford University.

- mhowes@stanford.edu

- `https://orcid.org/0009-0007-8048-3397`

- Department of Statistics
  Stanford University
  390 Jane Stanford Way
  Stanford, CA 94305-4020 USA

Austin van Loon is an assistant professor of work and organization studies at the MIT Sloan School of Management.

- vanloon@mit.edu

- `https://orcid.org/0000-0002-9565-7392`

- MIT Sloan School of Management
  100 Main St
  Cambridge, MA 02142 USA

Informed Consent

This study did not involve human subjects.

Data Availability Statement

A replication package with code and data will be available at `https://github.com/davidbroska/MixedSubjects`.