Leveraging AI & LLMs for Richer Insights into Public Opinion: Applications and Opportunities

Dr. Zachary P Dickson

z.dickson@lse.ac.uk
https://z-dickson.github.io/
Research Fellow in Quantitative Methods
Research Affiliate at the LSE Data Science Institute

May 2025

Language Models and Public Opinion Research

- LLMs can't replace a high quality opinion survey (*yet*)
  - Representativeness
  - Variance in responses
- LLMs are a powerful complement to measuring opinions and behaviour
- Exciting pathways for LLMs in survey research:
  - Dynamic/adaptive surveys and interviews
  - Generate/simulate new survey data*

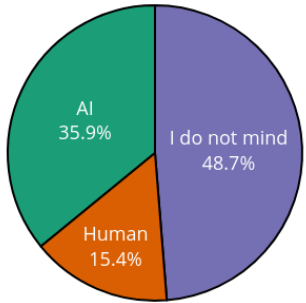
LLMs in [Adaptive] Survey Research

- Dynamic/adaptive surveys with embedded LLMs (Velez and Liu 2024)
  - Survey questions can be translated or modified to reflect user input
- Conversational surveys via interview (Geiecke and Jaravel 2024)
  - AI-led qualitative interviews at scale

"AI-powered interviewers can create a bridge between the richness of qualitative data and the statistical power of quantitative data" (Geicke & Jaravel, 2024)

Language Models in Interviewing for Electoral Choices

In the future, would you rather be interviewed by:



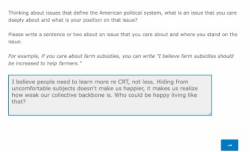
Source: Geicke & Jaravel (2024)

Dynamic/Adaptive Surveys

Velez and Liu (2024)

FIGURE 1. Dynamic Tailored Conjoint Design

(a) Open-ended question



(b) Personalized Conjoint

	Candidate 1	Candidate 2
Partisanship	Democrat	Democrat
Race/Ethnicity	Asian American	Black
Asylum Policy	Allow political refugees to obtain asylum	Allow economic refugees to obtain asylum
Religion	Catholic	Protestantism
Career	State representative	Agriculture
Veteran Status	Served	Did Not Serve
Critical Race Theory	The amount of education on Critical Race Theory should remain the same	People should learn more about Critical Race Theory, not less
Deportation Policy	Identify and deport all undocumented immigrants	Identify and deport all undocumented immigrants
Social Spending	Decrease spending on social programs that assist immigrants	Keep spending on social programs that assist immigrants the same
Sex	Female	Male
Border Wall Policy	Build a physical wall along the entire southern border	Build a physical wall along parts of the southern border
Age	62	66

Note: Screenshots of the survey design. (a) Participants begin by responding to an open-ended question about their most important issue, which is used to generate a personalized conjoint task. (b) Participants complete a set of candidate choice trials with their tailored issue position featured alongside immigration issue positions and demographic characteristics. See Appendix A for more details on conjoint task.

Synthetic Data Generation with LLMs

Examples:

- Simulated human samples (Argyle et al 2023)
  - Questionable reliability (Bisbee et al 2023)
- Behavioural game theory applications (Li et al 2023; Akata et al 2023)
- Simulated social media environment (using LLMs to interact w/ respondents) (Allamong et al 2024)
- LLM performance in wargames (Lamparth et al 2024)
- Do humans or LLMs make more convincing political arguments? (Palmer & Spirling 2024)

Simulating Public Opinion with LLMs

Bisbee et al (2024)

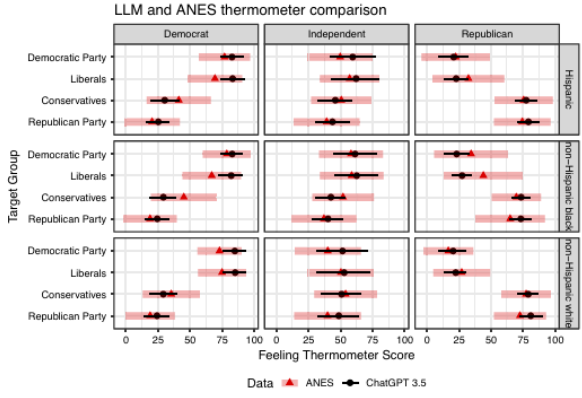
> "It is [YEAR]. You are a [AGE] year-old, [MARST], [RACETH] [GENDER] with [EDUCATION] making [INCOME] per year, living in the United States. You are [IDEO], [REGIS] [PID] who [INTEREST] pays attention to what's going on in government and politics."[13]

In each query to ChatGPT, the characteristics in brackets are substituted with values corresponding to a real respondent in the 2016 or 2020 wave of the ANES. These values include:

- [YEAR]: 2016 or 2020
- [AGE]: age in years of ANES respondent
- [RACETH]: non-Hispanic white, non-Hispanic black, or Hispanic
- [GENDER]: male or female
- [MARST]: divorced, married, separated, single, or widowed
- [EDUCATION]: a high school diploma, some college but no degree, a bachelor's degree or more
- [INCOME]: $30k, $50k, $80k, $100k, $150k or more
- [IDEO]: an extremely liberal, a liberal, a slightly liberal, a moderate, a slightly conservative, a conservative, an extremely conservative
- [REGIS]: registered, unregistered
- [PID]: Democrat, Independent, Republican
- [INTEREST]: never, sometimes, frequently, regularly, always

Can LLMs Represent Public Opinion?

Bisbee et al (2024)



The Path Forward: Opportunities & Responsible Innovation

- Opportunities
  - New insights from combining diverse data
    - Adaptive and dynamic surveys
    - Synthetic data*
    - Improving existing surveys (wording, cultural sensitivity etc.)
- Key Challenges & Considerations
  - Ethics: Privacy, potential for misuse (e.g., microtargeting, manipulation – Reddit story)
  - Validation: very early days for LLMs – be curious, but be caution and consult a range of experts outside your domain

Thank you for your attention!

- Slides available on: Github or via the QR code
- My email: z.dickson@lse.ac.uk

Enhancing Surveys with LLMs

- **Complementing existing survey methods**
  - Improving question wording (Wang et al. 2024)
  - Questionnaire pretesting in cross-cultural settings (Adhikari et al. 2025)
  - Analyzing open-ended questions/responses (Mellon et al. 2024)