

Exploring Agent-based Interviewing in Web Surveys

Jan Karem Höhne (DZHW, Leibniz University Hannover)

Marco Angrisani (University of Southern California)

Fred Conrad (University of Michigan)

Arie Kapteyn (University of Southern California)

Florian Keusch (University of Mannheim)

Session proposal

Web surveys continue to replace other survey modes, especially in-person interviews. Even large-scale surveys, such as the European Social Survey and the National Longitudinal Study of Adolescent to Adult Health, now routinely collect data via web surveys. However, the absence of interviewers complicates the provision of assistance to respondents and the creation of trust and motivation. This absence raises concern about answer quality. The advent of Generative Artificial Intelligence makes it possible to build interviewing agents that are visually realistic and conversationally responsive, deriving the latter ability from Large Language Models. Embedding such agents in web surveys promises to restore some of the quality-enhancing contributions of human interviews. Intelligent agents can clarify questions and provide feedback beyond what is typical in text-based web surveys and their mere presence can reduce speeding and non-differentiation but may introduce social desirability. Because respondents can choose an agent this may foster rapport helping to overcome social desirability. These innovations do not only give web surveys a human touch but make them more inclusive. Individuals with low literacy and education or who are not skilled speakers and/or readers of the survey language (e.g., immigrants and refugees) may be more likely to participate if they see (or hear) an agent that looks (or sounds) like them. Similarly, those with sensory challenges, especially the elderly, may favor verbal communication with a realistic looking, conversational agent, over text-based communication. In this session, we invite studies on all kinds of interviewing agents, not just those we have described here. This can be in various settings (lab or field) and with different study designs (cross-sectional or longitudinal). Contributions on legal and ethical considerations when using agent-based interviewing are also welcome. This similarly applies to studies that are work in progress.

Keywords: Answer behavior, Data quality, Embodied Agents, Large Language Models, Web surveys