Enhancing Open-Ended Survey Responses through LLM-Based Contextual Probing: An Experimental Study

[1]Mr.Dominik Wieland, [2]Mr. Nicolas Leyh

Technical University of Munich, TUM School of Management, Germany

Abstract

Probing - defined as follow-up questioning to encourage response elaboration - is an effective technique in qualitative research to improve the quality of open-ended responses. While its effectiveness in human-led interviews is established, it remains unclear whether these benefits transfer to AI-driven survey environments. Especially recent advances in large language models (LLMs) now enable chatbots to generate dynamic, context-sensitive probes in real time. This study investigates whether LLM-based contextual probing improves response quality in open-ended survey questions compared to rule-based and non-probing chatbot conditions. An experiment was conducted with 151 participants randomly assigned to one of three groups: no probing (EG1), rule-based non-contextual probing (EG2), and LLM- based contextual probing (EG3). All participants answered the same open-ended questions regarding a planned digital retirement savings platform. The chatbot in EG3 generated personalized follow-up questions using GPT-4o, while EG2 used a fixed generic probe. Data quality was assessed via two dimensions: response length (total word count) and thematic richness (number of distinct topics). The results show that contextual probing increased both response length and thematic richness compared to the other conditions. Statistical tests confirmed robust and significant differences between all groups, with the strongest effects observed in the LLM-enabled contextual probing group. These findings demonstrate that LLMs can replicate and enhance known probing effects from human interviewing, offering a scalable approach for improving qualitative data collection in digital surveys. The study contributes to the growing literature on human-AI interaction and provides practical implications for the design of conversational survey tools.

Keywords: Contextual probing; chatbot interaction; large language models; response quality