Position paper accepted for CONVERSATIONS 2024 – the 8th International Workshop on Chatbots and Human-Centred AI, hosted by CERTH, Thessaloniki, Greece, December 4-5, 2024.

Exploring Agent-Based Conversational Strategies for Enhancing Qualitative Survey Automation

Yunxing Liu[1 [0000−1111−2222−3333]] and Jean-Bernard Martens[1 [0000−0002−4414−1572]]

Eindhoven University of Technology , De Zaale 5600MB, Netherlands

Abstract. Advancements in artificial intelligence (AI) and large language models (LLMs), such as GPT-4, have opened new horizons for automating qualitative surveys, promising scalability and consistency in data collection. This paper explores the integration of agent-based conversational strategies into intelligent agents to enhance the automation of qualitative surveys. By leveraging LLM's capabilities, we aim to investigate how specific conversational strategies, such as positive framing, construct elaboration, contrast and reflection, and praise, can improve the richness and depth of data collected, as well as user engagement and satisfaction. The Repertory Grid Technique is employed as the methodological framework for this study, serving as a structured approach to elicit personal constructs from participants. We present a detailed research design that combines qualitative and quantitative data collection and analysis, utilizing a mixed-methods approach. Our findings have the potential to advance qualitative research methodologies by demonstrating how intelligent agents equipped with specialized conversational strategies can effectively automate complex qualitative tasks, contributing valuable insights to the field of conversation design.

Keywords: Qualitative Surveys· Conversational Strategies for chatbot· Conversation Agents· Crowdsourcing

1 Introduction

Qualitative research methods are indispensable to explore complex human behaviors, attitudes, and cognitive structures. The Repertory Grid Technique (RGT) is particularly effective in this domain, offering a structured framework to elicit personal constructs by comparing and contrasting different elements such as products, services, or experiences [18]. However, traditional RGT surveys are heavily based on human interviewers, introducing variability and potential bias that challenge scalability and consistency in large-scale qualitative research [17].

Advancements in artificial intelligence (AI) and large language models (LLMs), such as GPT-4 [3], have revolutionized the potential for automating qualitative surveys, providing scalable, consistent, and efficient data collection capabilities [21]. Intelligent agents with human-like conversational abilities offer a

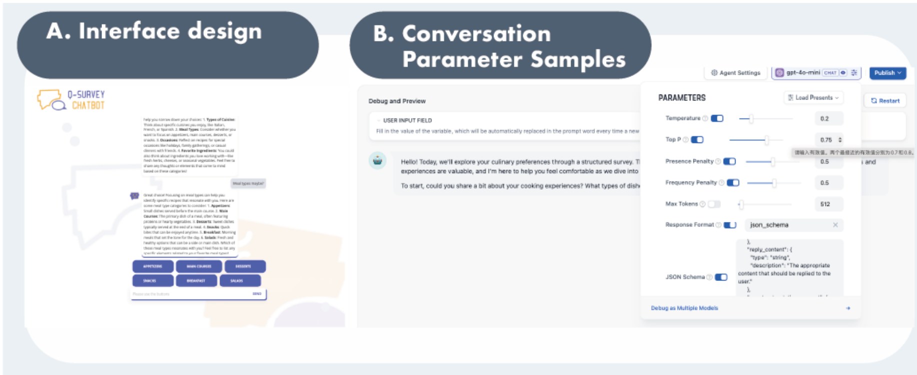

Fig. 1. This image provides an overview of the interface and technology used in this experiment. The A part includes buttons and the standard chatbot interface featuring the sample conversational interface. The B Part delineates the configuration of the chat agent and the application of the output schema from Dify [2].

unique opportunity to address the limitations inherent in traditional RGT methods, which typically rely on human interviewers. This reliance not only limits scalability but also introduces variability and potential bias, impacting the consistency and quality of collected data [17].

The adoption of chatbots in qualitative research, particularly in Repertory Grid Technique (RGT)-based studies, offers several benefits for large-scale, high-quality data collection. Chatbots enable researchers to overcome geographical and logistical barriers, reaching a broader participant pool in qualitative surveys[12]. They ensure consistent delivery of questions and management of responses, reducing variability found in traditional interviewer-led methods. However, balancing scalability with ethical and experiential considerations is important. Designing chatbot-driven surveys should prioritize user privacy, data security, and engagement to avoid shallow responses. Addressing these aspects allows for efficient automation while respecting participants' experiences.

In our previous work, we introduced the Q-Survey system, a hybrid user interface (HUI) combining both GUI and CUI components to facilitate qualitative surveys [19, 20]. While the HUI approach engaged users effectively, limitations in the conversational agent's abilities were noted due to the NLP and AI techniques available at that time. Participants reported that conversations could become tedious, which negatively impacted engagement, as indicated by low hedonic motivation scores in the UTAUT assessment [20]. The pilot study further highlighted a need for more sophisticated conversation design and enhanced agent capabilities.

This paper proposes enhancing the CUI component within the Q-Survey system by integrating intelligent agents equipped with specialized conversational strategies, harnessing the full capabilities of advanced LLMs. By focusing on a pure CUI approach and excluding GUI elements, we aim to enhance the richness and depth of construct elicitation in RGT-based qualitative surveys. This method addresses previous limitations and contributes to the field of human-computer interaction (HCI) by providing insights into the application of advanced AI usage in qualitative research.

2 Related Work

2.1 Conversational Agents in Qualitative Research

Conversational agents have gained attention in qualitative research for their potential to facilitate interactive and scalable data collection [15, 11]. Chatbots can simulate human-like interactions, making participants more comfortable and willing to share detailed information.

However, challenges persist in eliciting in-depth qualitative data comparable to human-led interviews due to limitations in natural language understanding. Hill et al. [16] found that while chatbots can handle simple queries, they struggle with complex or ambiguous responses, affecting data quality.

In our previous work [19, 20], we developed the Q-Survey system, integrating graphical and conversational elements for qualitative surveys. While effective in engaging users, limitations in the conversational agent's capabilities impacted the depth of data collected. Participants desired more natural and engaging interactions, highlighting the need for advanced conversational strategies to enhance qualitative survey automation.

2.2 Impact of Conversational Strategies

Conversational strategies are pivotal in determining the quality of interactions between users and conversational agents. Techniques such as positive framing, construct elaboration, contrast and reflection, and praise have been demonstrated to augment user engagement and enrich the data collected [7, 24]. Bickmore and Picard [8] identified that relational agents utilizing empathetic and encouraging language can form stronger bonds with users, thereby fostering increased trust and more open communication. The literature consistently indicates that research related to conversational agents using social dialogue can facilitate more engaging interactions, culminating in higher user satisfaction [13]. Furthermore, conversational strategies like social cues [14] and approaches to prevent conversational breakdowns [4, 6] have also been explored.

Despite these findings, a notable gap exists in the literature concerning the systematic integration of these strategies into intelligent agents for qualitative research. The majority of existing studies concentrate on task-oriented dialogues or customer service applications [21], with scant exploration of how these strategies affect qualitative data collection. Our previous studies [20] indicated that integrating conversational elements into surveys could enhance user engagement but also underscored the need for more sophisticated conversational designs. Participants reported that the conversations occasionally felt unnatural or repetitive, suggesting a requirement for the refinement of the agent's conversational strategies. Our current study aims to address this gap by providing empirical evidence on the integration of specific conversational strategies into intelligent agents to improve data quality and user engagement in qualitative surveys. We examine the impact of these strategies on construct elicitation (when participants compare and contrast elements to reveal their personal constructs—mental frameworks that define their perspectives on the elements in question [18])within the context of the RGT by utilizing advanced conversational agent techniques [23], thereby contributing to the evolution of more effective tools for the automation of qualitative research.

3 Research Design

The research presents a foundational framework to explore how conversational strategies can enhance the interactions of intelligent agents in RGT-based qualitative surveys. Serving as an initial proof of concept, the study aims to lay the groundwork for future automation of qualitative survey methodologies through intelligent conversational agents. By integrating specific strategies, it examines their effectiveness in enhancing data richness and construct elicitation using a mixed-method approach. The research focuses on the utility of conversational agents in HUI, particularly within the RGT framework[18], providing empirical insights and validating agent-led interactions for detailed data collection. The findings contribute to establishing a methodological approach adaptable for future studies.

3.1 Participants

We will recruit a diverse sample of participants aged 18 and above, ensuring variability across factors such as cooking experience, age, gender, cultural background, and culinary expertise. The purpose of this diversity is to enhance the generalizability of the findings and ensure that the insights gained from the study are applicable across different demographic groups.

To ensure sufficient statistical power and validity of results, we will recruit a minimum of 30 participants for each experimental condition, including four groups corresponding to the conversational strategies and one control group. In total, this will result in a minimum of 150 participants.

Participants will be recruited through Appen[1], a global platform that provides access to diverse participant pools for research purposes. The Appen's participant pool includes a wide range of individuals across various demographics, ensuring that we can recruit a balanced and diverse sample that meets the study's requirements.

3.2 Experimental Conditions

Participants will be randomly assigned to one of the following groups:

– **Experimental Group:** Agent employing all conversational strategies (Positive Framing, Construct Elaboration, Contrast and Reflection, Praise, and No strategy as a controlled group).
– **Control Group:** Agent without additional conversational strategies, using a standard conversational approach.

3.3 Procedures

1. **Introduction:**
   – Participants receive a brief overview of the study and provide informed consent on the Appen site.
   – They are instructed on how to interact with the intelligent agent, named **Q-Survey bot**, and informed that the topic is recipes.
2. **Interaction with the Intelligent Agent:**
   – Participants engage in a conversational survey with **Q-Survey bot**, following the guidelines and live hints specified in the agent's design.
   – The agent guides them to identify at least six elements (recipes) and assists them in eliciting at least five pairs of constructs using the triads method.
3. **Data Collection:**
   – Conversation logs are recorded for qualitative analysis.
   – Participants complete the Bot Usability Scale (BUS-15) questionnaire [9, 10], and a tailored Unified Theory of Acceptance and Use of Technology (UTAUT) questionnaire [22] to assess their perception of the system's usability and their acceptance of the technology.
   – An additional questionnaire is administered to evaluate the quality of the constructs generated.

3.4 Agent Design Details

The intelligent agent, **Q-survey bot**, is developed using GPT-4o mini and designed to incorporate several specialized conversational strategies aimed at enhancing user engagement and the quality of elicited data:

– **Positive Framing Strategy:** This strategy emphasizes the positive aspects of the participant's experience, encouraging them to approach responses with a constructive mindset and share richer details.
– **Construct Elaboration Strategy:** Encourages participants to delve deeper into their responses, facilitating more nuanced and comprehensive expressions of personal constructs.
– **Contrast and Reflection Strategy:** Guides participants to compare different constructs, helping them explore and articulate the relationships and distinctions between constructs.
– **Praise Strategy:** Provides positive reinforcement following effective responses, building participant confidence and promoting more enthusiastic engagement.

– **No Strategy (Control Group):** In this condition, the agent offers only essential suggestions or clarifications when necessary, allowing us to observe the effects of different strategies against a baseline approach.

The "recipes" theme for repertory grid research was chosen to align with previous studies, enabling consistency for comparative analyses while leveraging an accessible, universal topic. Recipes are easy to understand, allowing for broad participation across diverse cultural backgrounds, age groups, and culinary expertise levels. This topic also facilitates deeper exploration of personal preferences and cognitive patterns in choices and comparisons, supporting the generalizability and richness of insights gathered in this study.

In designing Q-survey bot, we applied a structured approach using the triad method [5], where the agent presents three elements (e.g., three different recipes) to the participant and prompts them to identify how one element differs from the others. This method effectively guides participants in reflective comparison, helping them recognize and articulate unique constructs. Sample conversations demonstrating the specific application of each strategy in the context of the RGT-based survey are provided in Appendix C.

3.5 Measurement Methods

To validate our research claims, we employ both quantitative and qualitative measurement methods (details in AppendixA):

– **Bot Usability Scale (BUS-15):** Assesses the usability of the chatbot from the user's perspective.
– **Tailored UTAUT:** Measures user acceptance of the technology, including factors such as performance expectancy, effort expectancy, social influence, and facilitating conditions.
– **Construct Quality Assessment Questionnaire:** Evaluates the richness, originality, and clarity of the constructs generated.
– **Content Analysis of Conversation Logs:** Qualitative analysis to assess the depth and quality of the data collected.

4 Expected Contributions

This research aims to contribute to the fields of HCI and qualitative research methodologies by:

– Demonstrating the effectiveness of integrating multiple conversational strategies into intelligent agents for RGT-based surveys.
– Providing empirical evidence on how these strategies impact data quality, user engagement, and survey efficiency.
– Offering detailed insights into agent design for qualitative research applications.

5 Conclusion

By integrating specialized conversational strategies into an intelligent agent, we aim to enhance the construct elicitation process in RGT-based qualitative surveys. Leveraging GPT-4o's capabilities allows for more natural and engaging interactions, potentially leading to richer data and improved participation satisfaction. This research has the potential to advance qualitative research methodologies and provide valuable insights into the application of AI in human-computer interaction.

References

1. Appen is the trusted partner for innovative world-class ai applications. https://appen.com/, (Accessed on 09/12/2023)
2. Dify.ai · the innovation engine for generative ai applications. https://dify.ai/, (Accessed on 09/25/2024)
3. Achiam, J., Adler, S., Agarwal, S., Ahmad, L., Akkaya, I., Aleman, F.L., Almeida, D., Altenschmidt, J., Altman, S., Anadkat, S., et al.: Gpt-4 technical report. arXiv preprint arXiv:2303.08774 (2023)
4. Ashktorab, Z., Jain, M., Liao, Q.V., Weisz, J.D.: Resilient chatbots: Repair strategy preferences for conversational breakdowns. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. p. 1–12. CHI '19, Association for Computing Machinery, New York, NY, USA (2019). https://doi.org/10.1145/3290605.3300484, https://doi.org/10.1145/3290605.3300484
5. Bell, R., Jankowicz, A.: Repertory Grid Technique: A Guide for Practitioners. John Wiley Sons, Hoboken, NJ (2003)
6. Benner, D., Elshan, E., Schöbel, S., Janson, A.: What do you mean? a review on recovery strategies to overcome conversational breakdowns of conversational agents. In: ICIS (2021)
7. Benson, M.: Conversational futures in online surveys. Survey Methodology **27**, 245–261 (2020)
8. Bickmore, T., Picard, R.: Establishing the computer–patient working alliance in automated health behavior change interventions. In: Patient education and counseling. vol. 59, pp. 21–30. Elsevier (2005)
9. Borsci, S., Malizia, A., Schmettow, M., Van Der Velde, F., Tariverdiyeva, G., Balaji, D., Chamberlain, A.: The chatbot usability scale: the design and pilot of a usability scale for interaction with ai-based conversational agents. Personal and ubiquitous computing **26**, 95–119 (2022)
10. Borsci, S., Schmettow, M., Malizia, A., Chamberlain, A., Van Der Velde, F.: A confirmatory factorial analysis of the chatbot usability scale: a multilanguage validation. Personal and Ubiquitous Computing **27**(2), 317–330 (2023)
11. Brandtzaeg, P.B., Følstad, A.: Why people use chatbots. In: International Conference on Internet Science. pp. 377–392. Springer (2017). https://doi.org/10.1007/978-3-319-70284-1_30
12. Burr, V.: Qualitative data collection using digital repertory grids. Journal of Qualitative Research **19**(3), 56–72 (2022). https://doi.org/10.1080/19312458.2022.2345678

13. Clark, L., Pantidi, N., Cooney, O., Doyle, P., Garaialde, D., Edwards, J., Spillane, B., Gilmartin, E., Murad, C., Munteanu, C., et al.: What makes a good conversation? challenges in designing truly conversational agents. In: Proceedings of the 2019 CHI conference on human factors in computing systems. pp. 1–12 (2019)
14. Feine, J., Gnewuch, U., Morana, S., Maedche, A.: A taxonomy of social cues for conversational agents. International Journal of Human-Computer Studies **132**, 138–161 (2019). https://doi.org/https://doi.org/10.1016/j.ijhcs.2019.07.009, https://www.sciencedirect.com/science/article/pii/S1071581918305238
15. Følstad, A., Brandtzæg, P.B.: Chatbots and the new world of HCI. Interactions **25**(5), 38–42 (2018). https://doi.org/10.1145/3236673
16. Hill, J., Ford, W.R., Farreras, I.G.: Real conversations with artificial intelligence: A comparison between human–human online conversations and human–chatbot conversations. Computers in Human Behavior **49**, 245–250 (2015). https://doi.org/10.1016/j.chb.2015.02.026
17. Jankowicz, D.: The easy guide to repertory grids. John wiley & sons (2005)
18. Kelly, G.: Personal construct psychology. Nueva York: Norton (1955)
19. Liu, Y., Martens, J.B.: Conversation-based hybrid user interface for structured qualitative survey: A pilot study using repertory grid. In: CHI Conference on Human Factors in Computing Systems Extended Abstracts. pp. 1–8 (2022)
20. Liu, Y., Martens, J.B.: Conversation-based hybrid ui for the repertory grid technique: A lab experiment into automation of qualitative surveys. International Journal of Human-Computer Studies **184**, 103227 (2024)
21. Shum, H.Y., He, X.D., Li, D.: From eliza to xiaoice: Challenges and opportunities with social chatbots. Frontiers of Information Technology & Electronic Engineering **19**(1), 10–26 (2018)
22. Venkatesh, V., Thong, J.Y., Xu, X.: Unified theory of acceptance and use of technology: A synthesis and the road ahead. Journal of the association for Information Systems **17**(5), 328–376 (2016)
23. Wang, L., Ma, C., Feng, X., Zhang, Z., Yang, H., Zhang, J., Chen, Z., Tang, J., Chen, X., Lin, Y., et al.: A survey on large language model based autonomous agents. Frontiers of Computer Science **18**(6), 186345 (2024)
24. Xiao, Z., Zhou, M.X., Liao, Q.V., Mark, G., Chi, C., Chen, W., Yang, H.: Tell me about yourself: Using an ai-powered chatbot to conduct conversational surveys with open-ended questions. ACM Transactions on Computer-Human Interaction (TOCHI) **27**(3), 1–37 (jun 2020). https://doi.org/10.1145/3381804, https://doi.org/10.1145/3381804

A BUS-15 Questionnaire

The BUS-15 is a 15-item questionnaire designed to measure users' perception of chatbot usability across several factors. To unify our data format, participants rate each item on a seven-point Likert scale from -3 ('Strongly disagree') to 3 ('Strongly agree').

A.1 Factor 1: Perceived Accessibility to Chatbot Functions

1. The chatbot function was easily detectable.
2. It was easy to find the chatbot.

A.2 Factor 2: Perceived Quality of Chatbot Functions

1. Communicating with the chatbot was clear.
2. I was immediately made aware of what information the chatbot can give me.
3. The interaction with the chatbot felt like an ongoing conversation.
4. The chatbot was able to keep track of context.
5. The chatbot was able to make references to the website or service when appropriate.
6. The chatbot could handle situations in which the line of conversation was not clear.
7. The chatbot's responses were easy to understand.

A.3 Factor 3: Perceived Quality of Conversation and Information Provided

1. I find that the chatbot understands what I want and helps me achieve my goal.
2. The chatbot gives me the appropriate amount of information.
3. The chatbot only gives me the information I need.
4. I feel like the chatbot's responses were accurate.

A.4 Factor 4: Perceived Privacy and Security

1. I believe the chatbot informs me of any possible privacy issues.

A.5 Factor 5: Time Response

1. My waiting time for a response from the chatbot was short.

B Construct Quality Assessment Questionnaire

Participants will be asked to rate the following statements on a seven-point Likert scale from -3 ('Strongly disagree') to 3 ('Strongly agree'):

1. The constructs I provided were detailed and comprehensive.
2. I felt that I could express my thoughts clearly during the conversation.
3. The chatbot helped me think deeply about the differences between the recipes.
4. The constructs captured my personal perspectives effectively.
5. The conversation was engaging and encouraged me to elaborate on my ideas.

C Prompt Design for the Intelligent Agent

The intelligent agent, **Q-survey bot**, is designed following specific guidelines to ensure effective construct elicitation:

C.1 Agent Introduction and Guidelines

- **Greeting and Introduction:** The bot introduces itself and explains the purpose of the survey.
- **Selecting Elements:** Participants are guided to list at least six recipes, using open-ended questions and the Single Choice Generating Strategy when necessary.
- **Construct Elicitation Using Triads:** The bot presents three recipes at a time and asks the participant to identify one that is different from the others, explaining the reason.
- **Conversational Strategies:** The bot employs the following strategies to enhance interaction:
  - Positive Framing Strategy
  - Construct Elaboration Strategy
  - Contrast and Reflection Strategy
  - Praise Strategy
  - No Specific strategy as comparsion
- **Use of Single Choice Options:** When providing or suggesting options, the bot places them in the *single_ choice* section instead of in the conversation.
- **Formatting Guidelines:** The bot uses Markdown formatting to enhance readability, emphasizing important words in bold and purple color.

C.2 Example Interaction Snippet

- **Bot:** "Hello! I am **Q-survey bot**. Today, we'll explore your culinary preferences through a structured survey. There are no right or wrong answers—just your personal insights."
- **Bot:** "Can you think of at least six recipes that you particularly enjoy or frequently prepare?"
- **Participant:** [Provides recipes]
- **Bot:** "Great! Now, let's look at three of these recipes: Recipe A, Recipe B, and Recipe C. Which one of these is different from the others, and why?"
- **Participant:** [Provides response]
- **Bot:** "**Excellent!** Can you tell me more about that difference?"

D Ethical Considerations

- **Informed Consent:** Participants will be informed about the purpose of the study, the procedures involved, and their rights, including the right to withdraw at any time without penalty.
- **Data Privacy:** All data collected will be anonymized and stored securely in the server of Tu/e. Personal identifiers will be not collected to protect participant confidentiality.
- **Transparency:** Participants will be debriefed about the study's aims and methods after their participation.

- **Compliance:** The study will adhere to ethical guidelines as stipulated by the institutional review board (ERB).
- **Potential Risks and Mitigation:** We anticipate minimal risk to participants with the Openai Moderation for both output and input. Should any other discomfort arise, participants can withdraw at any time without consequence.