Voice or Text? Exploring Optimal GPT Interaction Strategies to Enhance Student Learning Outcomes

Background:

Among the many recent changes in education, two trends have been particularly transformative. The first is the widespread integration of Large Language Models (LLMs) such as GPT-4 into educational contexts, demonstrating substantial positive impacts on students' learning outcomes[1]. The second trend is the growing interest in voice AI technology across various disciplines, including computer science, psychology, and education[2-3].

Previously, students' interactions with GPTs were limited to text-based inputs. However, the rise of voice AI technology provides new avenues for student-AI communication. These converging trends have led students to increasingly consider using voice interaction when learning with GPT-based systems. Existing research indicates that when learners interact through voice technology, they tend to focus more on clearly expressing their thoughts, reflecting more deeply on how to accurately convey their intentions before vocalizing their queries[4]. Therefore, it is essential to explore the interactive effects of interaction mode (voice vs. typing) and question type (open-ended vs. closed-ended) on students' learning outcomes during GPT-assisted learning.

Additionally, based on these effects, this research plans to develop a tutoring system utilizing GPT-4, capable of classifying question types and automatically recommending optimal interaction modes, thereby enhancing student learning outcomes.

Research Questions:

RQ1: When students interact with GPT models, what is the interactive effect between interaction mode (voice vs. typing) and question type (open-ended vs. closed-ended) on their learning outcomes?

RQ2: How can GPT-based interaction data be used to develop predictive models aimed at improving students' learning outcomes?

Research Design:

RQ1 Design(Lab experiment):

This study will conduct a 2 (interaction modality: voice vs. type) × 2 (question type: open-ended vs. closed-ended) between-subjects design.

RQ1 Procedure:

Participants were instructed to ask GPT-4 open-ended (closed-ended) questions using voice (typing), according to their assigned experimental conditions. They will have up to 3 minutes per question and can ask follow-up queries based on GPT-4' s responses. Learning effectiveness will be objectively assessed through immediate post-task tests measuring knowledge comprehension and memory, as well as a delayed test conducted one week later to evaluate long-term retention. Assessing Learning Engagement through Facial Expression Recognition and Eye-Tracking Technology[5]. To capture subjective learning outcomes, participants will complete questionnaires after the task, evaluating perceived learning effectiveness, technology adoption, usability, self-efficacy, and learning satisfaction .

RQ1 Data Analysis:

A two-way ANOVA will be conducted to examine the interaction effects of modality (voice vs. typing) and question type (open-ended vs. closed-ended) on learning outcomes.

RQ2:

Based on the interaction data collected during the RQ1 experiment (including question type, interaction modality, keyword count, query length, and number of follow-up questions), machine learning algorithms (e.g. Support Vector Machine) will be employed to develop a predictive model for student learning outcomes. Specifically, the data will first be randomly divided into training and testing datasets, with cross-validation used for model training and validation. Feature importance analysis will then identify the key interaction factors significantly influencing learning outcomes. Finally, the predictive model will be utilized to recommend the optimal interaction modality (voice or text) for different question types, enabling real-time personalized interaction strategies to enhance students' learning efficiency and outcomes.

References:

[1] Lin J, Han Z, Thomas D R, et al. How can i get it right? using gpt to rephrase incorrect trainee responses[J]. International journal of artificial intelligence in education, 2024: 1-27.

[2] Arnold R, Tas S, Hildebrandt C, et al. Any sirious concerns yet? – an empirical analysis of voice assistants' impact on consumer behavior and assessment of emerging policy challenges[C]//An Empirical Analysis of Voice Assistants' Impact on Consumer Behavior and Assessment of Emerging Policy Challenges (July 25, 2019). TPRC47: The 47th Research Conference on Communication, Information and Internet Policy. 2019.

[3] Crestani F, Du H. Written versus spoken queries: A qualitative and quantitative comparative analysis[J]. Journal of the American Society for Information Science and Technology, 2006, 57(7): 881-890.

[4] Melumad S. Vocalizing search: How voice technologies alter consumer search processes and satisfaction[J]. Journal of Consumer Research, 2023, 50(3): 533-553.

[5] D' Mello S, Olney A, Williams C, et al. Gaze tutor: A gaze-reactive intelligent tutoring system[J]. International Journal of human-computer studies, 2012, 70(5): 377-398.