HCI Assignment 1 Paper Reading:
Enhancing UX Evaluation Through Collaboration with Conversational AI Assistants

Group Members: 2350989 Qizheng Zhang,
2351050 Ruichen Yang

March 23, 2025

1 Introduction

The paper explores *the Integration of Conversational AI Assistants into UX Evaluation*, focusing on proactive dialogue and timing. Usability testing is a crucial aspect of user experience (UX) evaluation, yet traditional methods are labor-intensive and require extensive video analysis. Recent advancements in AI have introduced conversational AI assistants (CAs) as potential tools to support UX evaluation. This study investigates the effects of proactive AI-generated suggestions appearing at different times—before, in sync with, and after usability problems—to determine optimal timing for AI-assisted usability analysis.

2 Related Work

Prior research has investigated AI-driven usability problem detection using machine learning, natural language processing (NLP), and sentiment analysis. While these automated methods offer efficiency, they often lack contextual understanding and require human oversight. Human-AI collaborative methods, particularly conversational assistants, provide an interactive approach. Previous studies suggest that proactive AI interactions can enhance user trust and efficiency, but the impact of AI intervention timing remains underexplored.

3 Research Method

This study employs a hybrid Wizard-of-Oz approach, integrating ChatGPT-generated usability problem suggestions combined with a human moderator answering impromptu questions. A within-subjects experiment was conducted with 24 UX evaluators analyzing three usability test videos (website, mobile app, and VR interface). Automatic suggestions were provided at different timing conditions :before, during, or after usability problems. Researchers measured:

(i) Analytic performance: Number of usability problems identified.

(ii) Subjective perceptions: Trust, efficiency, and preference for AI timing.

(iii) Responses to AI suggestions: Agreement, correction, clarification requests, or disagreement.

4 Experimental Results

Key findings indicate that the timing of AI suggestions did not significantly impact the number of identified usability problems. However, suggestions after usability problems were preferred by participants, improving trust and efficiency. Participants primarily responded to AI suggestions by agreeing (77.6%), correcting, or requesting clarifications. And ChatGPT successfully identified 14 usability problems, but missed 58.8% of the unique problems found by human evaluators, indicating the irreplaceable role of human expertise. Issues missed by AI were related to UI design, interaction-based difficulties, navigation, and user expectations, highlighting AI's limitations in video-based UX analysis.

5 Insights

The study provides key insights into human-AI collaboration in UX evaluation. Firstly, AI is a validation tool. Most UX evaluators preferred to first analyze usability issues independently and then use AI suggestions for confirmation, reinforcing the importance of human judgment. And for AI timing, it should be more personalized. Since different users have varying workflow preferences, future AI assistants should allow customizable timing of suggestions. There're also AI's limitations: AI-generated suggestions were often incomplete or lacked deep contextual understanding, reinforcing the need for multimodal AI systems that analyze both verbal and visual cues.

Last but not least, some participants viewed AI as a competitor rather than a collaborator. Future AI tools should be designed to foster cooperation rather than rivalry in usability analysis.

6 Conclusion

This study provides valuable insights into the optimal integration of conversational AI assistants in UX evaluation. While proactive AI suggestions can enhance trust and efficiency, human expertise remains indispensable. Future research should focus on refining AI's ability to analyze multimodal data, improving explanation depth, and providing more adaptive AI collaboration strategies.