Exploring a Multimodal Chatbot as a Facilitator in Therapeutic Art Activity

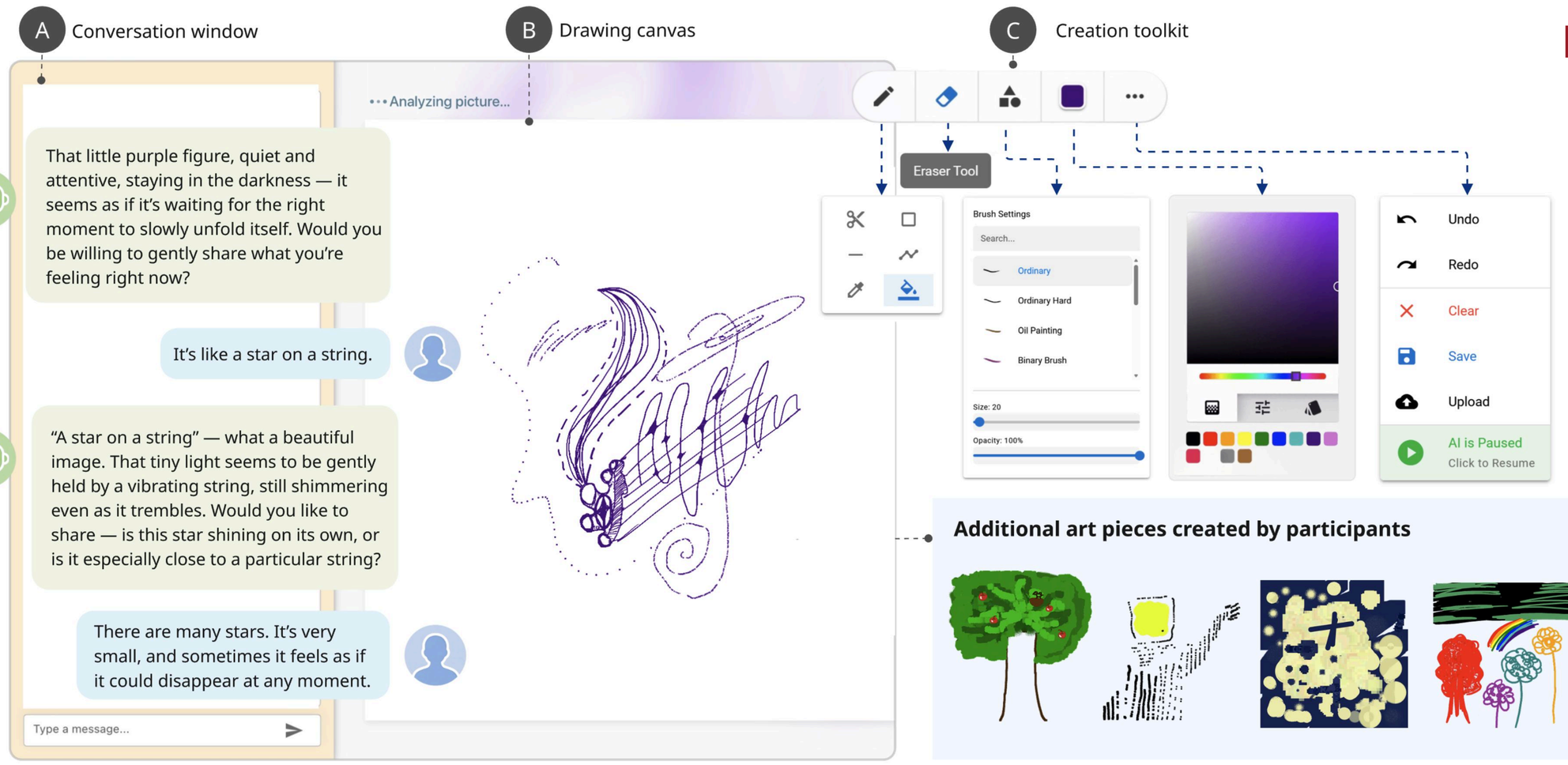
Le Lin[1], Zihao Zhu[1], Rainbow Tin Hung Ho[2], Jing Liao[1], Yuhan Luo[1]

香港城市大學 City University of Hong Kong

香 港 大 學 THE UNIVERSITY OF HONG KONG

DIS 2026

← Scan for Full Paper 👀

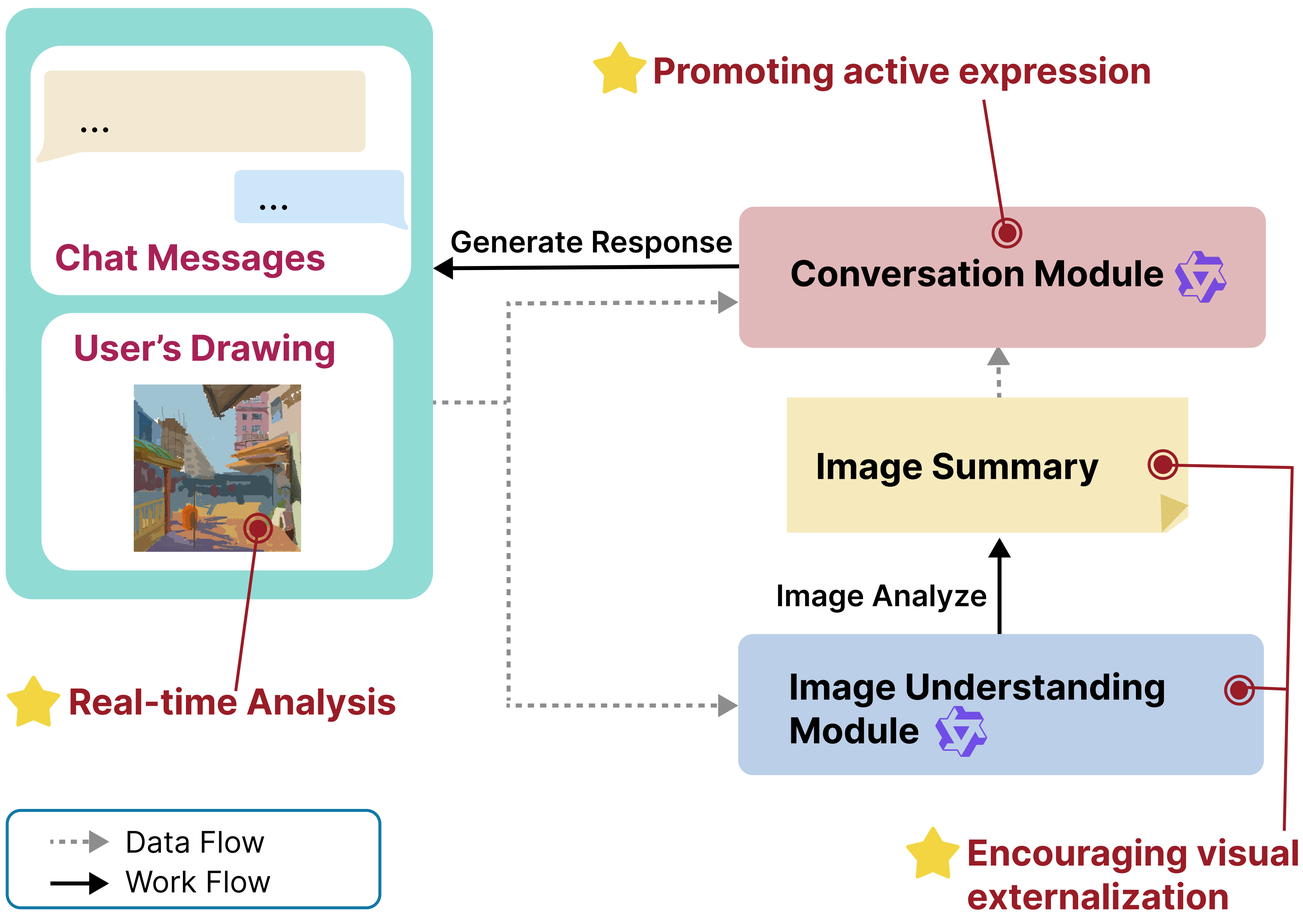


Introduction

We introduce a novel web-based, **MLLM-powered chatbot** featuring

- **real-time visual understanding**
- and **free-form drawing** capabilities,

bridging the gap between creative production and verbal reflection in **therapeutic art** activities.



Preliminary Findings

Key Benefits

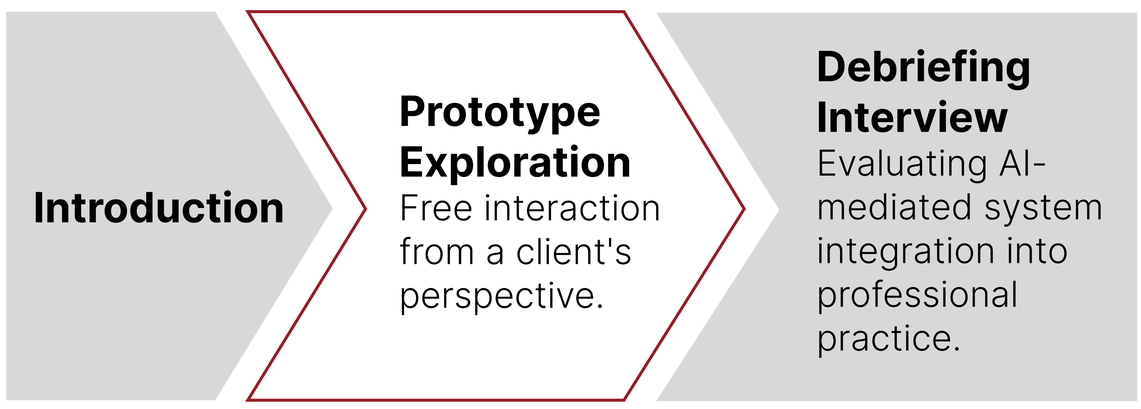
- **Low-stakes entry point** for emotional expression.
- Effective **visual comprehension**.
- **Supportive** therapeutic **dialogue**.

Recommended Improvements

- **Entryways & Risk Management.**
  Implement pre-session screening and monitor for therapeutic rupture.
- **Personalized Therapeutic Alignment**
  Tailoring sessions via user profiles, preferred styles, and long-term memory to build therapeutic alliance.
- **Balancing Therapeutic Depth and Width**
  Probing specific visual details (depth) while maintaining open-ended prompts to preserve imagination (width).
- **Enriching Visual Interactivity**
  Enrich art-making via contextual UI & human–AI co-creation.

Expert Evaluation

Participants: 5 experts in art therapy and related fields.



Future Work

- Optimize **conversation flows** to better balance interaction depth and width.
- Integrate on-canvas **visual features** for direct, multimodal AI-user communication.
- Develop a **risk management module** to monitor for signs of user distress.
- Conduct **longitudinal field studies** in **authentic contexts** using mixed-method evaluations.