



# The Synergy of Dialogue and Art: Exploring the Potential of Multimodal AI Chatbots in Emotional Support

Xueyang Wang

Tsinghua University

Zhongguancun Laboratory

Beijing, China

wang-xy22@mails.tsinghua.edu.cn

Runyan Tian

Tsinghua University

Beijing, China

try23@mails.tsinghua.edu.cn

Qiuyi Zeng

University of Cambridge

Cambridge, United Kingdom

qz299@cantab.ac.uk

Chenyue Tu

Chinese University of Hongkong,

Shenzhen

Shenzhen, China

chenyetu@link.cuhk.edu.cn

Shuning Zhang

Tsinghua University

Beijing, China

zhang-sn19@mails.tsinghua.edu.cn

Xin Yi\*

Tsinghua University

Zhongguancun Laboratory

Beijing, China

yixin@tsinghua.edu.cn

Hewu Li

Tsinghua University

Zhongguancun Laboratory

Beijing, China

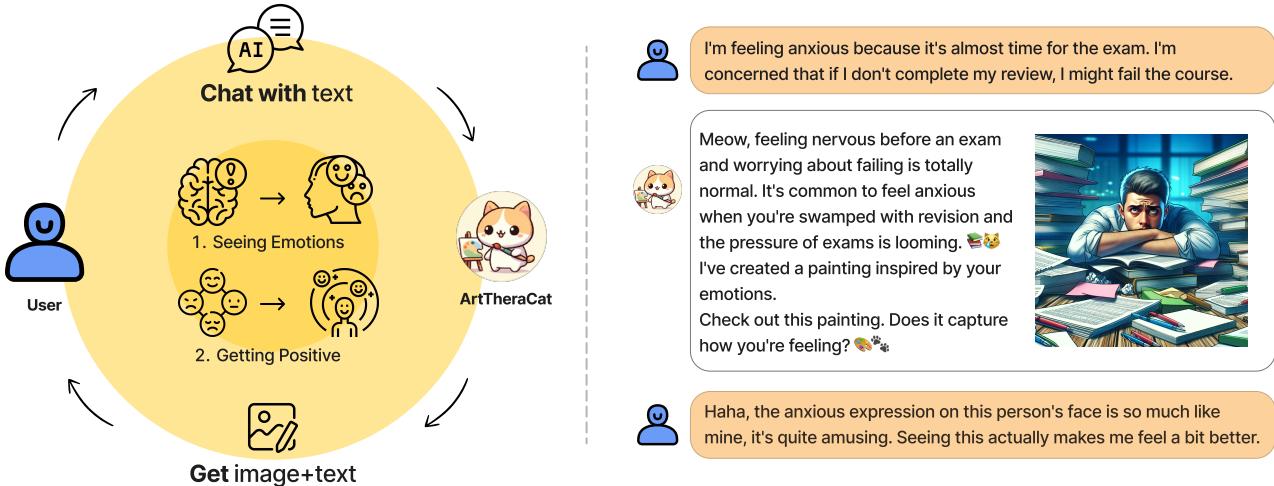
lihewu@cernet.edu.cn

Pei-Luen Patrick Rau

Tsinghua University

Beijing, China

rpl@tsinghua.edu.cn



**Figure 1:** We introduce ArtTheraCat, an innovative chatbot designed to support mental health through a unique fusion of AI-mediated dialogue and image generation. The left half of the figure illustrates our core idea: in conversation, ArtTheraCat aids users in visualizing their negative emotions through image generation, supplemented by textual soothing and guidance, to help improve their mood. The right half of the figure presents an example of a conversational round.

\*Corresponding author.



This work is licensed under a Creative Commons Attribution International 4.0 License.

## ABSTRACT

The rapid advancements in generative AI have spurred the development of AI chatbots for emotional support. In this work, we designed ArtTheraCat, a novel multimodal chatbot that promotes mental health by integrating supportive dialogue with artistic image generation. Utilizing images as an auxiliary medium, ArtTheraCat guides users through a process of emotion externalization in multi-turn dialogues, which can facilitate self-awareness, reflection, and

positive emotional encouragement. Results from a user study involving 21 participants indicate a significant improvement in their emotional state post-interaction with ArtTheraCat, especially in reducing negative affects. Qualitative analysis highlighted the role of generated images as a non-verbal medium in emotional release and self-insight. The findings of this paper provide practical guidance for the design of multimodal mental health chatbots.

## CCS CONCEPTS

- Human-centered computing → Empirical studies in interaction design; Empirical studies in HCI.

## KEYWORDS

Emotional Support, Multimodal Chatbot, Digital Art Therapy, Emotional Externalization, Affective Computing

### ACM Reference Format:

Xueyang Wang, Runyan Tian, Qiuyi Zeng, Chenye Tu, Shuning Zhang, Xin Yi, Hewu Li, and Pei-Luen Patrick Rau. 2024. The Synergy of Dialogue and Art: Exploring the Potential of Multimodal AI Chatbots in Emotional Support. In *Companion of the 2024 Computer-Supported Cooperative Work and Social Computing (CSCW Companion '24)*, November 9–13, 2024, San Jose, Costa Rica. ACM, New York, NY, USA, 7 pages. <https://doi.org/10.1145/3678884.3681843>

## 1 INTRODUCTION

The prevalence of mental health challenges is rising. Data from Mental Health America in 2023 indicates that about 21% of U.S. adults, or roughly 50 million people, suffer from at least one mental illness [46]. Yet, there's a global shortfall in psychotherapists, creating a significant gap in service provision [6]. This gap underscores the urgency of innovative approaches in mental support. One such approach is the use of AI-driven conversational agents [20, 34]. Current AI chatbots facilitate empathetic dialogue, providing a non-judgmental space for individuals to share emotions and thoughts [11, 39]. They can identify mental health issues early [43] and offer initial emotional support and guidance [19, 42].

In the realm of modern psychotherapy, art therapy, a unique discipline [32], employs artistic tools like paint, chalk, and markers, augmenting traditional conversational methods. This form of non-verbal communication is particularly effective for individuals who find it challenging to express their emotions verbally. Studies indicate that art-making and appreciation significantly aid in mood regulation [8, 33]. The act of creating art provides a secure avenue for expressing or transforming emotions, thus addressing deep-seated conflicts, fears, or sorrow. Additionally, evidence suggests that static visual art and music enable observers to perceive emotions, thereby enhancing empathy and emotional connection [10, 35], which contributes to emotional healing [13].

At present, digital art-making tools allow users to express their emotional states through digital artistry with AI acting as a supportive intermediary, identifying two primary applications: creative facilitation and analytical assistance. An example of creative facilitation is DeepThink [9], which uses AI brushes to render lifelike scenes based on user inputs. For analytical assistance, alphDAPR [24] helps psychologists by quantifying elements within artwork. Furthermore, the exploration of Virtual Reality (VR) in art therapy

has provided insights into its efficacy, offering immersive environments for supportive engagement [17, 40]. Despite these advancements, existing digital art therapy tools have yet to harness the full potential of generative AI models. Their complexity in artistic creation and lack of nuanced comprehension capabilities limit their ability to foster profound, emotionally resonant interactions and deliver empathetic support.

The rapid advancements in generative AI are epitomized by multimodal large language models such as GPT-4<sup>1</sup>, which integrate image generation capabilities from DALL-E 3<sup>2</sup> into their text-based conversational framework, allowing users to easily create fine images through text prompts. These developments present novel opportunities in emotional support, particularly in enriching supportive conversations with contextually generated visual art. However, the design of chatbots that integrate image generation with text for emotional support, and the effectiveness of this multimodal approach, have not been sufficiently studied or validated. To address this research gap, our research focuses on two main research questions:

- **RQ1: Can integrating text dialogue with image art generation in a multimodal chatbot provide effective emotional support?**
- **RQ2: How do text-based dialogue and image art individually contribute to emotional improvement?**

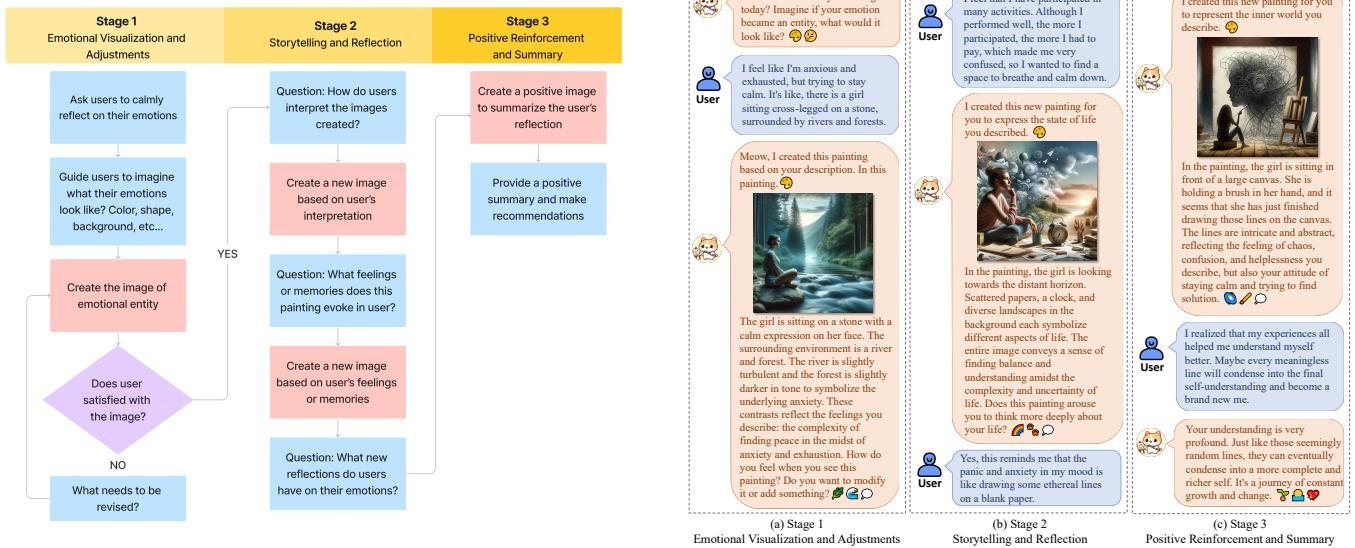
To address our research questions, we introduced ArtTheraCat, a novel mental health chatbot that combines conversational interaction with image creation, personified as a friendly, supportive cat. Mirroring traditional emotional externalization techniques in art therapy [32], ArtTheraCat's design is purposefully crafted to aid users in articulating and exploring their emotions through both artistic expression by text-to-image generation and reflective dialogues, aiming to reduce emotional distress. ArtTheraCat leverages the sophisticated natural language processing of GPT-4.0 along with the creative prowess of DALL-E 3's image generation. As illustrated in Figure 1, users engage with ArtTheraCat by describing their emotions in text, prompting the chatbot to generate visual representations of these emotions. Concurrently, it provides textual support and guidance.

In our preliminary study, we engaged 21 participants, all undergraduate and graduate students, during their final exam period—a time typically associated with heightened stress and anxiety [36]. Our quantitative analysis demonstrated a notable enhancement in the emotional state of participants post-interaction with ArtTheraCat. Specifically, there was a significant reduction in 9 out of the 10 negative affects measured by PANAS scale [7]. The qualitative analysis, drawn from user interviews, highlighted a synergistic effect from the amalgamation of chat and image art generation. Users appreciated the aesthetic appeal, rich associations, and deeper feelings prompted by the visual art, acknowledging the significant impact of non-verbal communication.

This study addresses several key gaps in current research on AI-driven emotional support and contributes to the CSCW field in multiple ways. Firstly, while existing studies have explored text-based AI chatbots or standalone digital art therapy tools, our work is

<sup>1</sup><https://openai.com/research/gpt-4>

<sup>2</sup><https://openai.com/dall-e-3>



**Figure 2: Flowchart illustrating the guiding process between ArtTheraCat and users. The blue process box represents dialogue guidance, and the red process box represents image generation.**

among the first to integrate these modalities, leveraging the synergy between verbal and visual communication in emotional support. Secondly, our findings provide insights into how AI can facilitate not just communication, but also emotional exploration and self-reflection. Lastly, this research contributes to the ongoing discourse in CSCW about the design of empathetic and engaging digital interfaces, offering practical guidelines for creating AI-driven tools that can foster trust and openness in users dealing with emotional challenges.

## 2 DESIGN OF ARTTHERACAT

### 2.1 Guiding Process: Emotional Exploration through Generated Art

ArtTheraCat integrates conversational AI with artistic image generation to provide emotional support. Its development was informed by extensive research on art therapy [21], emotionally expressive reflection [12, 45], and consultations with art therapists and clinical psychologists. The interaction process, illustrated in Figure 2, comprises three phases:

**1. Emotional Visualization and Adjustments:** Users are guided to identify and visualize their emotions as tangible objects, describing their characteristics. ArtTheraCat then uses DALL-E 3 to generate corresponding images, facilitating emotional externalization and understanding [22, 35, 47]. If needed, the image is collaboratively refined, mimicking the iterative nature of traditional art therapy [9, 32].

**2. Storytelling and Reflection:** Users discuss the generated artwork's relevance to their emotions, sharing thoughts and personal stories. This process aims to deepen emotional exploration and self-awareness [13, 14, 35]. Additional artworks may be created



**Figure 3: Illustration of a user journey through three stages. Excerpts from conversations between real users and ArtTheraCat.**

to reflect the evolving emotional narrative, supporting ongoing exploration and emotional expression [8, 33].

**3. Summary and Positive Reinforcement:** ArtTheraCat generates a final image with a positive tone, visually affirming the user's progress. It then summarizes the emotional journey, offering insights and suggestions for future emotional management. This closure aims to provide a sense of accomplishment and empower users for continued emotional growth [3, 16, 26, 27, 38].

### 2.2 User Interface and Implementation

ArtTheraCat was implemented using OpenAI's GPTs platform<sup>3</sup>, which integrates GPT-4 for natural language processing and DALL-E 3 for image generation. This implementation approach allows for a seamless integration of conversational AI and image creation capabilities without the need for a separate backend system. An example of ArtTheraCat's interaction with a user is illustrated in Figure 3, some representative conversations are excerpted.

The user interface is designed to be approachable and engaging. ArtTheraCat features a friendly cat avatar that "paints" the generated images. The interface uses a soft speaking tone, emoticons, and emojis to create a warm, whimsical atmosphere while maintaining a professional and sensitive approach to emotion interpretation and support [48].

The core functionality of ArtTheraCat is defined through carefully crafted instruction prompts within the GPTs interface. These prompts guide the AI to embody the persona of a supportive, art-therapy-oriented cat and outline the specific steps of the interaction process. Key features implemented through these prompts include:

- Emotion-to-image translation:** Instructions that guide the AI in interpreting users' emotional descriptions and converting them into suitable prompts for DALL-E 3.

<sup>3</sup><https://openai.com/blog/introducing-gpts>

- **Conversation stage management:** A structured approach that guides the AI through different stages of the conversation (e.g., initial visualization, reflection, positive reinforcement).
- **Image history tracking:** Instructions for the AI to maintain context and create thematically linked new images based on previous generations and user responses.

ArtTheraCat employs a flexible approach that adapts to various emotional states. The instruction prompts include guidelines for recognizing and responding to different emotions. For instance, when dealing with anger, the system is instructed to encourage more dynamic, expressive imagery and focus on constructive outlets. For sadness, it leans towards soothing colors and comforting metaphors. This adaptive approach ensures that each interaction is tailored to the user's specific emotional state, providing appropriate support and guidance.

The interaction's closure stage is initiated after the user has completed a predetermined series of reflection questions and image generations, typically occurring after 3-4 rounds of image creation and reflection. Users can also manually end the session at any point by expressing their desire to conclude the interaction, prompting ArtTheraCat to initiate the closure protocol with a final supportive message and session summary.

Prioritizing user privacy and ethical considerations, all interactions with ArtTheraCat are confidential and not shared with developers. The system's prompts are carefully configured to avoid creating distressing images, ensuring a safe and respectful environment for emotional exploration and support.

### 3 EVALUATION OF ARTTHERACAT

We carried out a preliminary user study evaluating ArtTheraCat's effectiveness. The study's goal was to examine user interactions and benefits derived from ArtTheraCat's AI-driven multi-modal dialogues, focusing on its influence on emotional support.

#### 3.1 Participants and Apparatus

The study involved 21 university students (14 females, 7 males; age:  $M = 22.7, SD = 2.2$ ) from diverse backgrounds, primarily from China (17), with others from Malaysia (2), Germany (1), and Canada (1). All participants were undergoing their final exam week, a period typically associated with high stress and negative emotions [36], providing an appropriate context for evaluating ArtTheraCat's impact.

Participants' familiarity with relevant technologies varied: 17 used chatbots at least weekly, 9 had experience with text-to-image tools, and 3 had previously used AI-based psychotherapy tools. Regarding psychotherapy familiarity, 13 had basic awareness, 6 showed moderate to high understanding, and 5 had attended psychotherapy sessions in the past year.

The study was conducted offline using an Asus laptop, with participants accessing ArtTheraCat's GPTs-based interface via Google Chrome. Data collection involved digital questionnaires and face-to-face interviews. Each session lasted 30-40 minutes, with participants receiving \$5 compensation. The study was approved by the university's ethics review board.

#### 3.2 Experiment Design and Procedure

Initially, participants completed a pre-experiment questionnaire in about 3 minutes, gathering basic demographics and employing the Positive and Negative Affect Schedule (PANAS) scale [7]. This scale, comprising two 10-item scales, rates affect on a 1 (Very slightly or not at all) to 5 (Extremely) scale.

Participants then engaged with ArtTheraCat for 10-15 minutes, experiencing its conversational and artistic generation features. We prioritized user privacy protection throughout the study. Participants were informed that their data would remain confidential and any research results would omit personal identifiers. During the ArtTheraCat interaction, researchers did not interfere or observe the session. Post-interaction, we sought consent again for data analysis; without consent, all related data, including dialogue records, were deleted.

Post-interaction, participants completed the PANAS scale again to measure their post-experiment emotional states. The experiment concluded with a 10-15 minute semi-structured interview, conducted verbally and audio-recorded. The interview explored participants' overall experiences with ArtTheraCat, focusing on three main areas: the interaction process, emotional improvement, and overall experience. Questions covered topics such as the fluency of communication, participants' interpretations of generated images, perceived changes in emotions or attitudes, and any concerns or suggestions for improvement. This approach provided rich qualitative insights to complement our quantitative findings from the PANAS scale, allowing for a comprehensive understanding of ArtTheraCat's impact on users' emotional states and experiences.

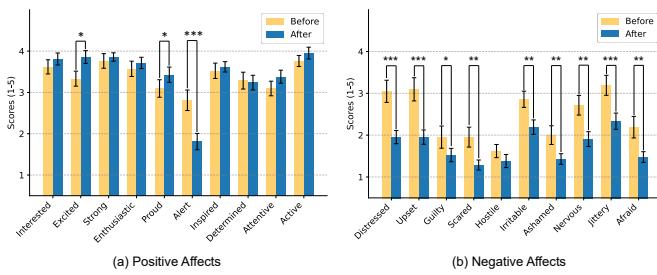
#### 3.3 Data Analysis

Our quantitative analysis focused on the variations in participants' emotional states before and after interacting with ArtTheraCat, as measured by the PANAS scale. After confirming the normality of the PANAS scale's affect differences pre- and post-interaction, we used paired t-tests to analyze changes in positive and negative affects.

The qualitative aspect of our study with ArtTheraCat utilized semi-structured interviews to delve into participants' experiences. We analyzed this qualitative data through thematic analysis [5]. Initially, four research team members ("coders") reviewed the interview transcripts for data immersion. Following this, an open coding process [23] was employed, with coders assigning relevant codes to specific data segments. Subsequent coder discussions focused on refining these codes, identifying and resolving overlaps, and discarding codes not pertinent to our research questions.

#### 3.4 Quantitative Results

As depicted in Figure 4(a), positive affects showed notable increases in **Excited** ( $\Delta Mean = 0.52, T = -2.14, p < .05$ ) and **Proud** ( $\Delta Mean = 0.33, T = -2.09, p < .05$ ), and a significant decrease in **Alert** ( $\Delta Mean = -1.0, T = 5.48, p < .001$ ). These results indicate that interacting with ArtTheraCat likely enhanced participants' excitement and pride, attributable to the engaging and innovative experience. The reduction in alertness may reflect a calming effect, reducing stress or anxiety.



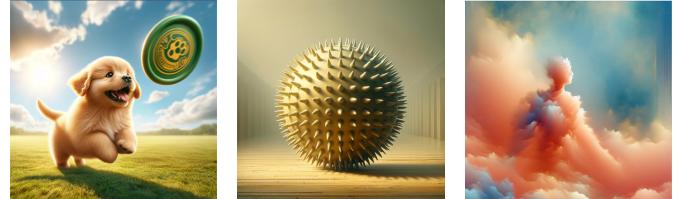
**Figure 4: Mean scores for positive (a) and negative (b) affective states pre- (orange) and post- (blue) interaction, measured on a Likert scale from 1 (Very slightly or not at all) to 5 (Extremely). Error bars denote standard errors. Statistical significance in paired samples is indicated by asterisks above the brackets linking the pre- and post-interaction means (\* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$ ).**

Regarding negative affects, as shown in Figure 4(b), significant reductions were observed across nine feelings, including **Distressed** ( $\Delta\text{Mean} = -1.1, T = 5.04, p < .001$ ), **Upset** ( $\Delta\text{Mean} = -1.14, T = 4.93, p < .001$ ), **Guilty** ( $\Delta\text{Mean} = -0.43, T = 2.26, p < .05$ ), **Scared** ( $\Delta\text{Mean} = -0.67, T = 3.16, p < .01$ ), **Irritable** ( $\Delta\text{Mean} = -0.67, T = 3.01, p < .01$ ), **Ashamed** ( $\Delta\text{Mean} = -0.57, T = 3.23, p < .01$ ), **Nervous** ( $\Delta\text{Mean} = -0.81, T = 3.30, p < .01$ ), **Jittery** ( $\Delta\text{Mean} = -0.86, T = 4.60, p < .001$ ), and **Afraid** ( $\Delta\text{Mean} = -0.71, T = 3.10, p < .01$ ).

### 3.5 Qualitative Results

**Emotional understanding through articulation and visualization:** Participants widely acknowledged that their interactions with ArtTheraCat served as a welcome pause from their busy routines. The process of expressing and depicting emotions facilitated a more profound comprehension and structuring of their internal states, a practice they rarely engaged in independently. This process was often described as tranquil and comforting. Participant P15 commented in the interview: “Initially, my emotional state was unclear, so my descriptions were sparse, leading to images that didn’t quite match my feelings. However, viewing these mismatched images spurred more emotional associations, which helped me swiftly sort out my true feelings.”

**Engagement and fulfillment in artistic interaction:** The participants expressed a deep sense of immersion through their interactions with ArtTheraCat. The process of fine-tuning the emotional representations led users into a state of flow [4], marked by heightened engagement and a rewarding sense of achievement once they received artwork that met their expectations. A notable example is Participant P6’s experience (Figure 5(a)), where the request for a specific image adjustment resulted in a depiction closely aligning with her envisioned concept. The user’s reaction, describing the final artwork as “super cute” and expressing happiness, underscores the gratification fostered through this creative process. Additionally, ArtTheraCat integrates users’ emotions and experiences into the generated images, enhancing their sense of connection with the artwork. As P12 noted in the interview, “It successfully incorporates



(a) Participant P6 expressed a desire to make the initial image of a puppy even cuter. In response, ArtTheraCat created a new image. P6 was delighted with the result, commenting, “Super cute! I like it, it perfectly matches what I had in mind.”

(b) In a state of mild anxiety, participant P10 described his emotion as a “yellow, spiky sphere in an empty space.” After viewing the visual representation, P10 remarked, “It doesn’t look as intimidating as I had imagined,” and indicated a reduction in his anxiety level.

(c) Participant P1 described her emotion as “peaceful, warm, and dreamy pale orange clouds.” P1 mentioned that the imagery brought back memories of watching clouds from an airplane, reminiscent of characters speaking to her from a fairy tale, evoking a sense of relaxation.

**Figure 5: Some real user cases in emotional visualization and adjustment process.**

*emotional expression into the drawing process, creating works that continuously evolve with my emotions or imagined scenes.”*

**Emotional relief through visualization:** The participants experienced significant relief from negative emotions by visualizing them as concrete objects. This approach helped transform abstract feelings into manageable forms. A case in point is Participant P10’s experience (Figure 5(b)). Describing his “somewhat anxious” state as a “yellow, spiky sphere in an empty space,” the participant noticed a reduction in the intensity of his anxiety upon seeing this visual representation. Further, this visualization led to an important insight in the response to ArtTheraCat: “Often, anxiety is an internal drain, a self-inflicted scare, which is entirely avoidable.”

**Enhanced emotional resonance through visual stimuli:** Users often observed that imagery had a more immediate and profound impact on their emotions than text. The visual representations not only conveyed feelings more directly but also triggered deeper emotional associations and resonances. This effect was particularly notable in how images stimulated memories and imagination, leading to relaxation and joy. For example, Participant P1’s experience with an image representing “peaceful, warm, and dreamy pale orange clouds” highlights this phenomenon (Figure 5(c)). This visualized emotion brought back comforting memories of “watching clouds from an airplane”, evoking a sense of nostalgia and fairy-tale-like wonder. Participant P19 emphasized the healing value of such visualizations in the interview, stating: “The greatest utility of emotional visualization is that it prompts me to recall those beautiful moments, which has a somewhat cure on my mood.”

**A fondness for the cat avatar:** The cat avatar and its friendly tone were highly praised for creating a relaxed, playful environment. This approach reduced the sense of formality, enhancing trust and openness in sharing emotions. P20 commented in the interview: “The paintbrush-wielding cat felt like a character from a fairy tale, bringing a magical, spirited quality.” Users also enjoyed the use of emojis, making the text more engaging and comfortable to read.

**Willingness to use in the future:** A significant proportion of participants (71.4%) expressed interest in continuing to use ArtTheraCat. They value its convenience for quickly addressing negative emotions, especially when access to mental health professionals is challenging. Those who were reluctant to continue using ArtTheraCat cited various reasons, including forgetting to use it, infrequent emotional issues, and a preference for voice interaction over typing.

## 4 DISCUSSIONS

### 4.1 RQ1: The Emotional Supportive Effect of Multimodal Chatbots

Our study reveals the significant potential of multimodal AI chatbots in providing emotional support, as evidenced by the marked improvements in users' emotional states following interaction with ArtTheraCat. The PANAS scale data showed notable increases in positive emotions coupled with substantial decreases in negative emotions, underscoring the efficacy of combining conversational AI with artistic image generation in creating a supportive digital environment.

ArtTheraCat's effectiveness stems from its multimodal approach to emotional support. At its core, the system guides users through a process of emotional visualization and adjustment, enabling them to translate their feelings into visual art. This unique form of emotional externalization, rooted in art therapy principles [31], facilitates deep introspection and leads to enhanced emotional clarity and catharsis [10, 44]. The interactive nature of this process, allowing users to refine their emotional representations, fosters a sense of agency crucial to the therapeutic process.

Building on this foundation, ArtTheraCat engages users in a stage of storytelling and reflection, leveraging the power of narrative in emotional processing. By encouraging users to construct stories around their visualized emotions, the system taps into the therapeutic potential of storytelling [2, 37]. This approach not only facilitates emotional expression but also promotes cognitive restructuring, allowing users to reframe their experiences in more adaptive ways. The subsequent positive reinforcement phase further solidifies these emotional gains, potentially contributing to longer-lasting positive effects.

Equally important is the engaging and safe interaction environment created by ArtTheraCat. Its design, featuring a friendly avatar and empathetic responses, addresses a critical challenge in digital mental health interventions - user engagement and trust [15]. By fostering a sense of safety and openness, ArtTheraCat increases user willingness to engage in meaningful emotional work, a crucial factor for therapeutic success.

### 4.2 RQ2: The Respective Roles of Dialogue and Image Generation

Our research reveals the profound impact of integrating textual dialogue with image generation in emotional support, demonstrating how these modalities complement and enhance each other to create a uniquely effective therapeutic experience.

The textual dialogue, powered by GPT-4's advanced language understanding, forms the foundation of ArtTheraCat's emotional support capability. It enables fluid, coherent, and empathetic conversations that adapt to users' emotional states. The AI's ability to interpret nuanced and fragmented expressions allows users to communicate their emotions freely, without the pressure of articulating complex feelings in detail [1]. This flexibility in communication is crucial for emotional exploration, as it mirrors the often non-linear nature of emotional processing.

Complementing this verbal interaction, the image generation component introduces a powerful non-verbal dimension to emotional support. Visual representations of emotions offer immediate and profound impact, eliciting rich associations that words alone might struggle to convey. The process of visualizing emotions as tangible images engages users in a form of artistic critique and appreciation [10], facilitating a deeper deconstruction of emotional meanings. This visual externalization makes abstract feelings more concrete and manageable, offering users new perspectives on their emotional landscapes.

The synergy between dialogue and image generation in ArtTheraCat creates a holistic and immersive experience that transcends the limitations of traditional single-modality approaches. While dialogue initiates and guides emotional reflection, image creation provides a canvas for emotional expression and exploration. This dual-modality approach allows users to engage with their emotions both verbally and visually, leading to a more comprehensive understanding and processing of their feelings [33].

## 5 LIMITATIONS AND FUTURE RESEARCH

Despite these promising results, our study has several limitations that point to directions for future research. The primary limitation is the narrow demographic of participants, predominantly students aged 20–25 years, and the absence of individuals diagnosed with mental disorders such as depression or anxiety, who are key potential users of such tools [43]. To address this, future studies will involve a broader user base and include professional psychotherapists and art therapists in the evaluation process [9]. This expansion will provide a more comprehensive understanding of ArtTheraCat's efficacy across diverse populations and clinical settings.

Another significant limitation is the lack of comparative experiments. While our preliminary tests demonstrated the effectiveness of multimodal chatbots in providing emotional support, comparative studies are necessary to evaluate this method against traditional text-based chatbots and other common emotional therapies. Such studies will help isolate the unique benefits of combining dialogue with image generation and provide insights into which aspects of ArtTheraCat contribute most significantly to its therapeutic effects. Additionally, the long-term assessment of ArtTheraCat's impact on mental health warrants attention. Future research could explore the chatbot's potential in recording and visualizing emotional changes [25, 28], similar to Apple Health's praised feature [30], offering users a more profound and aesthetically rich experience of their emotional journey over time.

Currently, ArtTheraCat's interaction is limited to text input, which may not fully capture the nuances of human emotional expression. Future research directions include integrating users' voice, expressions, and body actions as inputs for artistic creation [18], enhancing the chatbot's interactivity and potentially providing a more holistic emotional support experience. Furthermore, exploring additional supportive mediums such as music, video, and 3D generation could evoke varied responses from users, reflecting current trends in AI research [29, 41]. These enhancements could lead to more personalized and effective emotional support interventions, pushing the boundaries of AI-assisted mental health care.

## REFERENCES

- [1] AC Abbing, EW Baars, O Van Haastrecht, AS Ponstein, et al. 2019. Acceptance of anxiety through art therapy: a case report exploring how anthroposophic art therapy addresses emotion regulation and executive functioning. *Case reports in psychiatry* 2019 (2019).
- [2] Thomas D Carlson. 1997. Using art in narrative therapy: Enhancing therapeutic possibilities. *American Journal of Family Therapy* 25, 3 (1997), 271–283.
- [3] Charles S Carver, Michael F Scheier, and Suzanne C Segerstrom. 2010. Optimism. *Clinical psychology review* 30, 7 (2010), 879–889.
- [4] Tatiana Chemi. 2016. The experience of flow in artistic creation. *Flow experience: Empirical research and applications* (2016), 37–50.
- [5] Harris Ed Cooper, Paul M Camic, Debra L Long, AT Panter, David Ed Rindskopf, and Kenneth J Sher. 2012. *APA handbook of research methods in psychology, Vol 2: Research designs: Quantitative, qualitative, neuropsychological, and biological*. American Psychological Association.
- [6] Nathaniel Counts. 2023. Understanding the US Behavioral Health Workforce Shortage. *The Commonwealth Fund* (2023).
- [7] John R Crawford and Julie D Henry. 2004. The Positive and Negative Affect Schedule (PANAS): Construct validity, measurement properties and normative data in a large non-clinical sample. *British journal of clinical psychology* 43, 3 (2004), 245–265.
- [8] Lili De Petriollo and Ellen Winner. 2005. Does art improve mood? A test of a key assumption underlying art therapy. *Art Therapy* 22, 4 (2005), 205–212.
- [9] Xuejun Du, Pengcheng An, Justin Leung, April Li, Linda E Chapman, and Jian Zhao. 2024. DeepTHInk: Designing and probing human-AI co-creation in digital art therapy. *International Journal of Human-Computer Studies* 181 (2024), 103139.
- [10] Marina Ebert, Jessica D Hoffmann, Zorana Ivcevic, Christine Phan, and Marc A Brackett. 2015. Creativity, emotion, and art: Development and initial evaluation of a workshop for professional adults. *The International Journal of Creativity & Problem Solving* 25, 2 (2015), 47–59.
- [11] Amelia Fiske, Peter Henningsen, and Alena Buyx. 2019. Your robot therapist will see you now: ethical implications of embodied artificial intelligence in psychiatry, psychology, and psychotherapy. *Journal of medical Internet research* 21, 5 (2019), e13216.
- [12] Rowanne Fleck and Geraldine Fitzpatrick. 2010. Reflecting on reflection: framing a design landscape. In *Proceedings of the 22nd conference of the computer-human interaction special interest group of australia on computer-human interaction*. 216–223.
- [13] Zsuzsanna Geréb Valachiné, Adél Dancsik, Michelle M Fitos, and Renáta Cserjési. 2023. Online art therapy-based self-help intervention serving emotional betterment during COVID-19. *Art Therapy* 40, 3 (2023), 134–141.
- [14] LESLIE S Greenberg. 2002. Emotion-focused therapy: Coaching clients to work through their feelings. Washington, DC: American Psychological Association.
- Leichsenring, F., & Leibing, E.(2003). The effectiveness of psychodynamic therapy and cognitive behavior therapy in the treatment of personality disorders: A metaanalysis. *American Journal of Psychiatry* 160 (2002), 12231232.
- [15] Leslie S Greenberg. 2004. Emotion-focused therapy. *Clinical Psychology & Psychotherapy: An International Journal of Theory & Practice* 11, 1 (2004), 3–16.
- [16] JAMES J. GROSS. 2002. Emotion regulation: Affective, cognitive, and social consequences. *Psychophysiology* 39, 3 (2002), 281–291. <https://doi.org/10.1017/S0048577201393198>
- [17] Irit Hacmun, Dafna Regev, and Roy Salomon. 2021. Artistic creation in virtual reality for art therapy: A qualitative study with expert art therapists. *The Arts in Psychotherapy* 72 (2021), 101745.
- [18] Li Hu, Xin Gao, Peng Zhang, Ke Sun, Bang Zhang, and Liefeng Bo. 2023. Animate Anyone: Consistent and Controllable Image-to-Video Synthesis for Character Animation. *arXiv preprint arXiv:2311.17117* (2023).
- [19] Kevin W Jin, Qiwei Li, Yang Xie, and Guanhua Xiao. 2023. Artificial intelligence in mental healthcare: an overview and future perspectives. *The British Journal of Radiology* 96, 1150 (2023), 20230213.
- [20] Eunkyung Jo, Daniel A Epstein, Hyunhoon Jung, and Young-Ho Kim. 2023. Understanding the benefits and challenges of deploying conversational AI leveraging large language models for public health intervention. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–16.
- [21] Sandra L Kagan and Vija B Lusebrink. 1978. The expressive therapies continuum. *Art Psychotherapy* (1978).
- [22] Margaret L Keeling and Maria Bermudez. 2006. Externalizing problems through art and writing: Experience of process and helpfulness. *Journal of marital and family therapy* 32, 4 (2006), 405–419.
- [23] Shahedul Huq Khandkar. 2009. Open coding. *University of Calgary* 23, 2009 (2009).
- [24] Jiwon Kim, Jiwon Kang, Taeeun Kim, Hayeon Song, and Jinyoung Han. 2023. AlphaDAPR: An AI-based Explainable Expert Support System for Art Therapy. In *Proceedings of the 28th International Conference on Intelligent User Interfaces*. 19–31.
- [25] Taewan Kim, Seolyeong Bae, Hyun Ah Kim, Su-woo Lee, Hwajung Hong, Chanmo Yang, and Young-Ho Kim. 2024. MindfulDiary: Harnessing Large Language Model to Support Psychiatric Patients' Journaling. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–20.
- [26] Sander L Koole. 2009. The Psychology of Emotion Regulation: An Integrative Review. *Cognition and Emotion* 23, 1 (2009), 4–41. <https://doi.org/10.1080/02699930802619031>
- [27] Nichole Kryla-Lighthall and Mara Mather. 2009. The role of cognitive control in older adults' emotional well-being. (2009).
- [28] Kwangyoung Lee and Hwajung Hong. 2017. Designing for self-tracking of emotion and experience with tangible modality. In *Proceedings of the 2017 Conference on Designing Interactive Systems*. 465–475.
- [29] Chen-Hsuan Lin, Jun Gao, Luming Tang, Towsaki Takikawa, Xiaohui Zeng, Xun Huang, Karsten Kreis, Sanja Fidler, Ming-Yu Liu, and Tsung-Yi Lin. 2023. Magic3d: High-resolution text-to-3d content creation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 300–309.
- [30] Gough Yumu Lui, Dervla Loughnane, Caitlin Polley, Titus Jayarathna, and Paul P Breem. 2022. The apple watch for monitoring mental health-related physiological symptoms: Literature review. *JMIR Mental Health* 9, 9 (2022), e37354.
- [31] Cathy A Malchiodi. 2003. Art therapy and the brain. *Handbook of art therapy* (2003), 16–24.
- [32] Cathy A Malchiodi. 2011. *Handbook of art therapy*. Guilford Press.
- [33] Anastasia V Mishina, Juliya L Blinova, and Olga G Belomoyeva. 2017. Multimodal art therapy for overcoming negative emotional states among adolescents. *Helix* 8 (2017), 2307–11.
- [34] Surapon Nochaiwong, Chidchanok Ruengorn, Kednapa Thavorn, Brian Hutton, Ratanaporn Arwichan, Chabaphai Phosuya, Yongyuth Ruanta, Nahathai Wongpakaran, and Timakon Wongpakaran. 2021. Global prevalence of mental health issues among the general population during the coronavirus disease-2019 pandemic: a systematic review and meta-analysis. *Scientific reports* 11, 1 (2021), 10173.
- [35] Matthew Pelowski, Eva Specker, Gernot Gerger, Helmut Leder, and Lauren S Weingarten. 2020. Do you feel like I do? A study of spontaneous and deliberate emotion sharing and understanding between artists and perceivers of installation art. *Psychology of Aesthetics, Creativity, and the Arts* 14, 3 (2020), 276.
- [36] Adele Pitt, Florin Oprescu, Geraldine Tapia, and Marion Gray. 2018. An exploratory study of students' weekly stress levels and sources of stress during the semester. *Active Learning in Higher Education* 19, 1 (2018), 61–75.
- [37] Shirley Riley. 1997. Social constructionism: The narrative approach and clinical art therapy. *Art therapy* 14, 4 (1997), 282–284.
- [38] Martin EP Seligman. 2006. *Learned optimism: How to change your mind and your life*. Vintage.
- [39] Woosuk Seo, Chanmo Yang, and Young-Ho Kim. 2024. ChaCha: Leveraging Large Language Models to Prompt Children to Share Their Emotions about Personal Events. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–20.
- [40] Liat Shamri Zeevi. 2021. Making art therapy virtual: integrating virtual reality into art therapy with adolescents. *Frontiers in Psychology* 12 (2021), 584943.
- [41] Uriel Singer, Adam Polyak, Thomas Hayes, Xi Yin, Jie An, Songyang Zhang, Qiyuan Hu, Harry Yang, Oron Ashual, Oran Gafni, et al. 2022. Make-a-video: Text-to-video generation without text-video data. *arXiv preprint arXiv:2209.14792* (2022).
- [42] Inwha Song, Sachin R Pendse, Neha Kumar, and Munmun De Choudhury. 2024. The typing cure: Experiences with large language model chatbots for mental health support. *arXiv preprint arXiv:2401.14362* (2024).
- [43] Jie Sun, Qun-Xi Dong, San-Wang Wang, Yong-Bo Zheng, Xiao-Xing Liu, Tang-Sheng Lu, Kai Yuan, Jie Shi, Bin Hu, Lin Lu, et al. 2023. Artificial intelligence in psychiatry research, diagnosis, and therapy. *Asian Journal of Psychiatry* (2023), 103705.
- [44] Pablo PL Tinio. 2013. From artistic creation to aesthetic reception: The mirror model of art. *Psychology of Aesthetics, Creativity, and the Arts* 7, 3 (2013), 265.
- [45] Nadine Wagener, Leon Reicherts, Nima Zargham, Natalie Bartłomiejczyk, Ava Elizabeth Scott, Katherine Wang, Marit Bentvelzen, Evropi Stefanidi, Thomas Mildner, Yvonne Rogers, et al. 2023. SelVRReflect: A Guided VR Experience Fostering Reflection on Personal Challenges. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–17.
- [46] Darice Warren. [n. d.]. The State of Mental Health in America 2023: Adult Prevalence and Access to Care - NextStep Solutions – nssbehavioralhealth.com. <https://www.nssbehavioralhealth.com/nss-blog-the-state-of-mental-health-in-america-2023-adult-prevalence-and-access-to-care/>. [Accessed 16-01-2024].
- [47] Michael White and Michael Kingsley White. 2007. *Maps of narrative practice*. WW Norton & Company.
- [48] Mingrui Ray Zhang, Ruolin Wang, Xuhai Xu, Qisheng Li, Ather Sharif, and Jacob O Wobrock. 2021. Voicemoji: Emoji entry using voice for visually impaired people. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–18.