

In Dialogue with Poor Images: Exploring the Therapeutic Potential of Co-Creating Visual Art with AI

Sybil Liu(s.liu.16@umail.leidenuniv.nl)

Media Technology, Leiden University

Niels Bohrweg 1, 2333 CA Leiden

Supervisors: Rob Saunders (r.saunders@liacs.leidenuniv.nl) and Tessa Verhoef (t.verhoef@liacs.leidenuniv.nl)

Abstract

Building on the concept of creativity as an intrinsically rewarding activity, this study investigates whether and how co-creating art with AI demonstrates therapeutic potential, defined by mood regulation in therapeutic art-making (improved mood and flow) and activation of therapeutic factors in art therapy (self-expression and reflection). Conducted as an online Zoom workshop, the experiment involved 20 participants randomly assigned to either a text-to-image or sketch-to-image condition. The experiment included a non-directive phase and a task-oriented phase, with participants completing the Affect Grid and Flow Short Scale, followed by interviews. Results indicate that the text-to-image condition experienced significantly improved valence, arousal, and flow, while the sketch-to-image condition showed positive but less consistent effects. Thematic analysis of interviews highlighted participants' overall enjoyment and smooth experience, emphasizing the value of experimentation and gradual concretization. Based on these findings, the study suggests that depending on the therapeutic goal, different aspects of the interaction with AI should be emphasized. Balancing freedom and control and fostering dialogues between the participant, AI-generated art, and the AI tool can enhance mood and reflection respectively. Future research can explore how multimodal inputs and outputs, and varied task designs can tailor AI-assisted art-making to specific therapeutic contexts.

Introduction

Art captures things we have no words for. This capacity has been continually explored and expanded by generations of artists, weaving symbols, metaphors, emotions, and narratives into human creative history. Art-making, as a method employed by psychologists and therapists to probe into the less accessible internal worlds of humans, is a more recent endeavor. From the psychoanalytic theory of symbolism to imagery rescripting in cognitive behavior therapy (Hogan, 2016), images take a central place on the canvas of the unconscious mind, making it possible to express and re-imagine oneself through creative means.

The process of surfacing subjective experiences through visual methods, can be found in many applications beyond. Drawing, collage, photography and other similar participatory forms (Lierat, 2013; Suprpto et al., 2020) have proven valuable in qualitative studies across various target groups, extending the space for understanding and empathy.

With the rapid spread and iteration of AI image generation tools, the myth and magic of creativity seem to have been demystified, as people without art training can quickly create images that amaze viewers at first glance. Several studies have already explored AI as a creative assistant for professionals, potentially augmenting human creativity and productivity in terms of research and ideation (Lin et al., 2020), while its application in therapeutic practices remains less examined. This research seeks to address this gap by investigating the therapeutic potential of co-creating art with generative AI, specifically focusing on mood regulation and the activation of therapeutic factors.

Related Work

Image-Making in Art Therapy vs. Therapeutic Art-Making

Creative Arts Therapies (CATs) encompass a range of creative activities including, but not limited to, art, dance, drama, and music therapy, which are gaining increasing research attention for their positive psychological and physiological outcomes (de Witte et al., 2021). This study will focus on visual art-making as the rising popularity of image-making AI tools suggests opportunities and questions for their uses in therapeutic practices. It is, however, important to first make a distinction between art therapy and therapeutic art-making, which incorporate similar elements but carry different implications. Art therapy typically involves guided sessions with a trained therapist who helps individuals explore emotions, thoughts, and experiences through artistic expression. In contrast, therapeutic art-making is more flexible in offering distraction or entertainment as it can occur without the

presence of a therapist and take place in a variety of settings. While this distinction is useful when considering their implications, one should note that “[a]rt making in mental health settings...sit[s] along a practice continuum [which] includes: individual art practice, the open studio approach, community arts, health promotion orientations, art education, and art therapy” (Van Lith, Fenner, & Schofield, 2010). In practice, their different approaches can be mixed, tailored, and applied to meet individual needs.

Therapeutic Art-Making

Studies have explored the mood-elevating effects of art-making, focusing on how different media and strategies impact emotional regulation. After inducing sadness in participants, researchers found that drawing led to significantly more positive mood changes than writing, and distraction (drawing an unrelated or positive image) was a more effective strategy than venting (drawing an image related to the negative event) (Drake et al., 2011 & 2012). Both studies confirm that visual art-making can enhance mood, particularly when used as a distraction from negative feelings. Therapeutic art-making in general demonstrates effectiveness in emotion regulation in the short term, with upregulation of positive affect and distraction strategy as the more significant effect and strategy (Gruber & Oepen, 2018). A more recent study complements these findings with a proposed theoretical model identifying high flow, high arousal, and low rumination as the essential factors for mood-repair in personalized art-making activities; specifically, positive mood is enhanced by activities that reduce self-focus and promote a growth-oriented mindset (Futterman, Collier & Wayment, 2021). In addition to mood improvement, many community-based participatory projects or individual-initiated art-making activities have been shown to promote wellbeing and quality of life for participants with physical or psychological afflictions (Kim, Loring, & Kwekkeboom, 2018; Stickley, Wright, & Slade, 2018), though involvement of longer-term, regular professional guidance and art therapy interventions is more common in this context.

Art Therapy Theories and Practices

In art therapy practices, images and image-making are used in multiple ways, with different theories and models explaining and validating a myriad of possible goals and mechanisms.

The Expressive Therapies Continuum (ETC) (Hinz, 2019) is a commonly used model for applying and assessing creative activities in therapy. Several HCI studies utilized this framework for developing and evaluating the therapeutic meaning of the tools they developed (Du et al., 2024; Liu et al., 2024). The ETC outlines four levels of experience in creative functioning: the Kinesthetic/Sensory focuses on physicality and bodily engagement; the

Perceptual/Affective level for organization visual elements and verbal or non-verbal expression of ideas and emotions; the Cognitive/Symbolic level involves rational thinking and the use of symbols and metaphors in art-making; and the Creative level integrates the previous three levels, demonstrating the fulfillment achieved through the overall creative experience (Hinz, 2019; Liu et al., 2024).

De Witte et al. (2021) proposed a framework of 19 domains of Creative Arts Therapies (CAT) to identify therapeutic factors that bring about positive changes. Three domains unique to CAT are highlighted: embodiment (body awareness, physicality), concretization (creating visible products to facilitate verbal reflection and self-knowledge), and symbolism and metaphors (unconscious self-expression). Specific to art therapy, the following therapeutic factors were identified: concretization (visual narrative of life, portray feelings of past/future); symbolism and metaphor (unconscious self-expression); modulating time and space (flow state); understanding (reflection on art / own patterns, self awareness through artwork/explore one’s feelings, insights in emotions).

Drawing insights from both art-making in short-term mood regulation and art therapy practices, while it is evident that certain specific factors, such as embodiment (engaging with physical art materials) and non-verbal expression, may be partially lost in AI-assisted image generation, co-creating with AI can potentially still offer therapeutic benefits in terms of facilitating short-term mood elevation through achieving enjoyment and flow experience, and improving self-understanding through extracting symbols and encouraging reflections. Therefore, it is helpful to investigate how AI can effectively support these therapeutic processes and what specific interactions are enhanced or diminished through its use.

Based upon the idea that creativity can be an autotelic activity which is intrinsically rewarding and enjoyable (Compton and Mateas, 2015), this study sets out to explore the therapeutic potential of co-creating art with text-guided image generation AI, where therapeutic potential is informed by and defined as 1. achieving short term mood regulation as demonstrated in therapeutic art making and 2. eliciting similar therapeutic mechanisms in art therapy.

Image Generative AI and Well-being

Several recent studies have begun to explore the intersection of image-generative AI tools and mental health. The existing research can be broadly categorized into three areas:

Evaluating the quality of generated images. User evaluations based on specified criteria are used to examine

if generated images bring satisfaction (Kwon et al., 2024), align with user intent and context, and facilitate expression of events and emotions (Lee et al., 2023);

Process-focused investigation of AI's role in therapy. By investigating how users interact with AI tools and AI-generated materials, studies evaluated the benefits and implications of incorporating AI in therapy settings. This often involves development of new tools with expert review and user evaluation testing their effectiveness. For example, Yoo et al. (2023) developed Mind Palette, an AI chatbot for AI-assisted art-making and appreciation. Wan et al. (2024) developed a system for providing metaphor suggestions and generating visual metaphors called Metamorpheus. Finally, Du et al. (2024) developed DeepThInk, a human-AI co-creative drawing tool.

Co-design Studies for AI tools in Therapy Contexts. Collaboration with professional therapists generated design considerations and highlighted the importance of incorporating professional knowledge, supporting multi-stakeholder interaction, providing personalized treatment content, and enhancing user engagement (Han and Cai, 2023; Du et al., 2024).

Observations from existing studies confirmed AI's role in lowering the expertise threshold for art-making and supporting users' creativity and meaning-making through levels of the ETC (Liu et al., 2024; Du et al., 2024), while raising concerns about AI's unpredictability and its potential to overshadow user participation, agency, and ownership (Wan et al., 2024). On the other hand, no study has yet focused on the immediate or short-term effect on user affect after engaging in image-making tasks with AI, whether verbal or textual expressions hinder the therapeutic process, or which specific interactions in human-AI co-creative tasks demonstrate therapeutic potential.

Casual Creators

Compton and Mateas (2015) defined a category of creativity tools, Casual Creators, as “an interactive system that encourages the fast, confident, and pleasurable exploration of a possibility space, resulting in the creation or discovery of surprising new artifacts that bring feelings of pride, ownership, and creativity to the users that make them” (2015). Designed to support autotelic creativity rather than productive task completion, a casual creator therefore highlights intrinsic motivation, enjoyment, and play. Usually these tools are designed to support the creation of a particular type of artifact and encourage the intrinsic enjoyment derived from the creative process. One example discussed by Compton and Mateas (2015) was “Let's Create! Pottery”, a virtual pottery making app where users can create customized pottery designs by selecting

from material, ornament, and color options provided. Aspects of casual creators can also be found in many mainstream games which, besides the core gameplay, allow players to enjoy creating unique avatars by customizing their characters' appearance, clothing, and accessories with a range of given options.

Two concepts are central to casual creators are flow and possibility space. According to Csikszentmihalyi's Flow Theory, the sense of progress, instant feedback, and the balance between perceived skills and challenges contribute to a flow state. The possibility space refers to all potential artifacts that could be created using a particular tool. A successful casual creator, according to Compton and Mateas (2015), allows users to confidently navigate this space by quickly manipulating the system and evaluating the effects—the flow state is achieved through users progressively searching for and discovering the good artifacts. Meanwhile, the possibility space can be narrow to “provide greater support for the user, eliminating potential bad artifacts and speeding the process of creating good ones, at the expense of flexibility and versatility” (Compton and Mateas, 2015).

Connecting AI, Casual Creators, and Therapeutic Practices

While many image generative AI tools are not specifically designed for autotelic creativity, and extensive discussion surrounds its usage for productivity support, the typical process of generating images with text-guided AI is similar to casual creators in several aspects. These AI tools offer a vast possibility space by providing users with considerable freedom to experiment with different prompts, styles, and parameters, resulting in the discovery of unexpected and unique visual outputs. In the creative process facilitated by AI, users can similarly engage in a cycle of rapid evaluation, modification, and reflection. With each generation, users assess and select from a limited range of generated outcomes and quickly generate variations for refining their inputs and subsequent outputs. One example that incorporates generative AI while retaining features of casual creators is DeepThInk, an AI-infused art-making tool developed for digital art therapy settings, which offers both AI Brush and User Brush and enable choices of Styling and Filtering to encourage user participation in limited possibility space (Du et al., 2024).

Some characteristics of common text-guided image generation tools also align, to varying extents, with the design patterns identified by Compton and Mateas (2015), such as “the chorus line” (generating and presenting multiple artifacts), “mutant shopping” (suggested alternatives), “no blank canvas” (prompt suggestions), and “saving and sharing”. These systems may be lacking in

other aspects, however, such as “instant feedback” (immediately visible changes), “limiting actions to encourage exploration” (limited number of choices), “entertaining evaluations” (having a simulated critic), and “modifying the meaningful” (specific modifications), either due to current technical challenges and unpredictability in AI generations, or additional skills required from non-expert users. Thus, this leaves opportunities for investigating what fosters or restricts user’s engagement with the AI system, especially when directed at therapeutic purposes.

Therapeutic Computational Creativity (TCC) involves systems designed to enhance well-being and potentially improve mental health. In "A Roadmap for Therapeutic Computational Creativity," Pease et al. (2022) noted that casual creators contribute to TCC by providing accessible and enjoyable creative experiences. However, they also suggested that TCC should also “extend beyond casual creators” by encouraging greater user input and incorporating creative artifacts into the healing process. This offers opportunities for establishing and clarifying how AI, casual creators, and art-making can inform and complement each other to support therapeutic purposes (fig.1). Autotelic creativity, which prioritizes the enjoyable experience of exploratory creativity over task completion, aligns with therapeutic art-making’s emphasis on process rather than outcome. The flow state, a crucial goal for casual creators, mirrors a key therapeutic factor in both art-making and art therapy. Additionally, the possibility space offered by image generative AI allows users to set self-defined goals, evaluate, and modify outcomes, potentially encouraging reflection and understanding, which are key components of art therapy. By zooming in on the process of engaging with AI, it is possible to uncover specific interactions that demonstrate therapeutic potential. Here, therapeutic potential is defined as achieving similar effects of therapeutic art-making, which are improved mood and flow experience, and activating therapeutic factors in art therapy, which include expression and reflection.

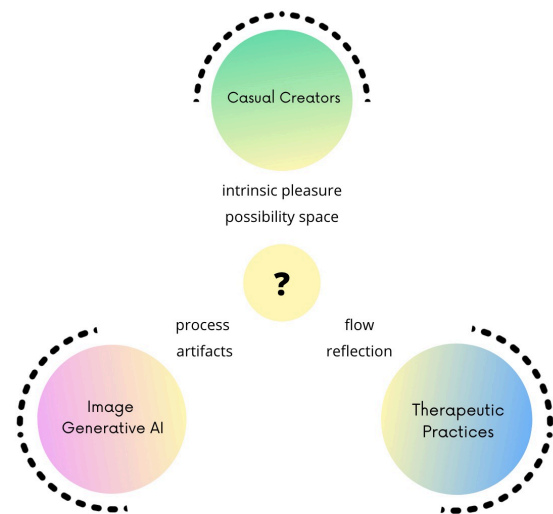


Figure 1: Connecting AI, Casual Creators, and Therapeutic Practices

Research Questions

The potential for image generative AI to support therapeutic outcomes through creative engagement requires investigation into the specific ways in which interaction with AI may influence these outcomes. Currently, text-to-image and sketch-to-image are two of the most commonly used strategies in AI-assisted art-making, while different therapeutic mechanisms may be at play in these two conditions. For instance, the text-to-image condition might rely more on cognitive and verbal processing, while the sketch-to-image condition allows participants to not only use text but also incorporate their own sketches into the AI-assisted creation process, preserving the embodied interaction in the traditional art-making process. They can start with expressing emotions and ideas nonverbally, capturing elements that may be difficult to capture through text alone. On the other hand, using AI to generate images from hand-drawn sketches entail different user experiences, and AI-generated images may not always capture the nuanced emotional or symbolic meanings intended by participants in their sketches. Therefore, this study aims to explore whether and how co-creating art with AI demonstrates therapeutic potential through the following research questions:

RQ1: Does co-creating visual art with AI demonstrate therapeutic potential through improved affect, facilitating flow experience, and achieving self-expression and reflection?

Hypothesis: Participants in both the text-to-image and sketch-to-image conditions will experience therapeutic effects, as evidenced by improved affect, achieving flow

experience, and encouraging self-expression and reflection. Specifically, individuals in the sketch-to-image condition will show more pronounced improvements for retaining embodiment and non-verbal expression.

The separation between conditions allows for an investigation into whether and how specific factors like the loss of embodiment and non-verbal expression impact the therapeutic effects. It is also possible to examine if the text-only condition compensates for these losses by offering other benefits. Therefore, by comparing the outcomes within and between the text- and sketch-to-image conditions, the study further evaluates whether and how different modes of interaction with AI impact therapeutic effects.

RQ2: How can co-creating with image generative AI support enjoyment, self-expression and reflection?

In traditional art therapy, this is done through extracting symbols and metaphors, facilitating self-understanding and emotion elicitation. RQ2 will focus on identifying similar opportunities and potential obstacles throughout the human-AI co-creation process.

Experiment

Set-up

The experiment took the form of an one-to-one online workshop session hosted on Zoom, where participants learned to use DreamStudio¹ through a provided video tutorial and create images following given instructions. The Zoom meeting (audio and screen sharing) was recorded to document their interactions with the tool and the follow-up interview.

Common (commercial) text-guided image generation tools include DALL-E 2, MidJourney, Stable Diffusion that use natural language descriptions to create images. These tools range from producing highly stylized art (MidJourney) to open-source customization (Stable Diffusion), catering to various creative and practical applications. For this experiment, DreamStudio, a text-to-image web application developed by Stability AI using Stable Diffusion model, is selected considering its user-friendly interface, multi-modal input methods, prompt and style suggestions, and the option to generate variations to facilitate exploration and refinement.

Participants were randomly assigned to one of two conditions: text-to-image or sketch-to-image. In the text-to-image condition, they used text as the only input for image generation, while in the sketch-to-image condition,

participants made sketches on paper, uploaded images of their sketch as input, and generated images guided by text prompts.

Observations From Pilot Tests

Prior to the experiments, two pilot tests were conducted, one text-to-image and one sketch-to-image, running through the workshop structure and measures to evaluate the feasibility and effectiveness of the procedure. Both participants are Master's students. Based on their feedback, several changes were implemented to ensure the smooth flow of the experiment:

Measurement Tools: Alternatives for quantitatively measuring the effects, such as the Positive and Negative Affect Schedule (PANAS, Watson et al., 1988) and the General Self-Efficacy Scale (Schwarzer and Jerusalem, 1995) were considered, but replaced with the Affect Grid (Russell et al., 1989) and the Flow State Scale (Rheinberg et al., 2003) to avoid lengthy question items and better capture participants' states at the moment, focusing on the immediate effects of the experience.

Clarity of Instructions: The tutorial and instructions were improved for clarity. Examples illustrating how participants can create both literal representations and use symbols and metaphors to express their experiences were added.

Language Accessibility: Since many participants were non-native English speakers, the instructions emphasized that text prompts do not need to be complete or grammatically perfect sentences. Participants were encouraged to focus on the image in their minds rather than on the accuracy of the text

Comparison of Approaches: The participant in the sketch-to-image condition highlighted a comparison between different approaches to using the drawing: as a brainstorming tool to uncover hidden ideas versus using photographs or online images to enhance the quality of the output. This will be further discussed.

Online Experiment Procedure

Participants first learn to use Dreamstudio by watching a short video tutorial, after which they are given some time to try out the tool and ask questions, with access to a quick guide throughout the workshop. Before the experiment, a pre-task survey gathered demographic data and included questions on participants' prior experience with similar image generation tools and their previous experience with art-making. During the experiment, participants shared their screen and their interactions with the tool were recorded, including the follow-up interview.

The experiment consists of two parts:

¹ <https://dreamstudio.ai/generate>

Part 1: Non-directive: Participants will have 10 to 15 minutes to create any images they choose, without any specific guidance or constraints.

Part 2: Task-oriented: Participants will engage in specific image creation tasks for 15 to 20 minutes, including: Self-Portraits: Past, Present, and Future; Stressor and Solution: Produce an image of a stressor and then create images depicting ways to manage or alleviate this stress; Nightmare Reimagined: Draw a past nightmare and then create images that alter or reimagine it in a preferred way. Participants are encouraged to use both literal representations and symbolic elements to convey their ideas, emotions, and experiences. After each part, participants are instructed to pick and share two to three of their favorite images.

Part 1 relates to therapeutic art-making by allowing participants to engage in free art-making without direction or moderation from facilitators. The goal is to examine if the process of engaging with image generative AI demonstrates any effect on mood regulation and flow state. Part 2 aligns more closely with art therapy tasks, where participants are given prompts to explore and express personal experiences and emotions. Specifically, the task options took inspiration from cognitive behavioral art therapy (CBAT), where art-making can be used to depict and manipulate mental images, reshaping negative thinking to alternative, adaptive meanings (Rosal & Gussak, 2015; Hogan, 2016). Part 2 thus aims to explore opportunities and limitations of co-creating images with AI when applied in art therapy context. However, it should be noted that the comparison of effectiveness between goal-oriented instructions and non-directive approaches for emotion regulation in specific contexts remains an ongoing discussion (Rankanen, 2016). In this exploratory study, it was hypothesized that Part 1 would bring greater mood improvement and flow experience due to the freedom it offers, while the structured tasks in Part 2 might be more effective in facilitating reflection and expression.

Measurements

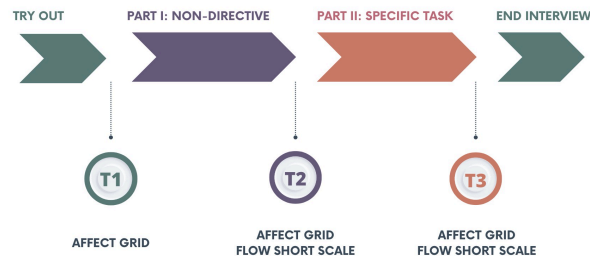


Figure 2: Experiment Procedure and Measurement Timings

The Affect Grid (Russell et al., 1989) was used to assess the participants’ valence and arousal before using the tool (T1), after finishing Part 1 (T2), and after finishing Part 2 (T3).

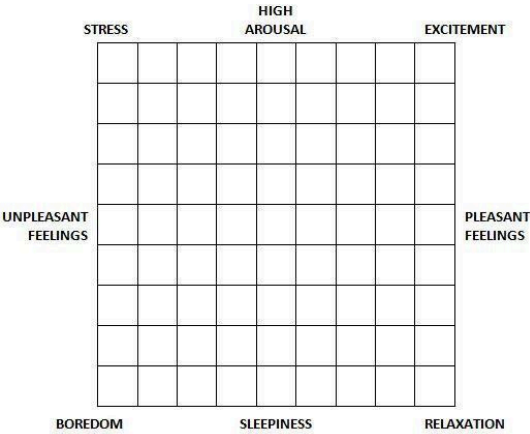


Figure 3: Example of Affect Grid

Previous studies on the effects of art-making on short-term mood regulation and emotional responses have employed the Affect Grid to assess participants’ affective state, confirming its utility for quick and repeated assessments (Dalebroux et al., 2008). In this grid of 9*9 empty boxes, participants were told to click in a single box to indicate how they were feeling at that moment. Scores along each dimension range from 1 to 9, horizontally from unpleasant to pleasant (valence) and vertically from sleepiness to excitement (arousal). The valence and arousal score were subsequently determined by locating the box checked.

Participants finished the Flow Short Scale (Rheinberg et al., 2003) after Part 1 (T2), and after Part 2 (T3), to measure the flow perceived by the participants after engaging with the tasks.

The scale includes 10 item²s evaluating aspects of the flow experience using a seven-point Likert scale from "strongly disagree" to "strongly agree." The total Flow score is determined by the sum of these items. Additionally, two subscores Fluency (items 2, 4, 5, 7, 8, 9) and Absorption (items 1, 3, 6, 10) can be calculated. These items involve perceptions about challenge, control, and concentration when engaged in this co-creative activity, which are key therapeutic factors explored in art therapy.

² The original Flow Short Scale also includes 3 worry items, but the 10 flow items can be used separately.

Flow-Items:

- 1 I feel just the right amount of challenge.
- 2 My thoughts/activities run fluidly and smoothly.
- 3 I don't notice time passing.
- 4 I have no difficulty concentrating.
- 5 My mind is completely clear.
- 6 I am totally absorbed in what I am doing.
- 7 The right thoughts/movements occur of their own accord.
- 8 I know what I have to do each step of the way.
- 9 I feel that I have everything under control.
- 10 I am completely lost in thought.

Figure 3: Flow Short Scale Items

Upon finishing the experiment, semi-structured interviews are conducted to review the processes with participants. The goal of the interview is to identify potential therapeutic qualities of co-creating with AI in terms of therapeutic factors such as concretization, symbolism, and understanding, as well as identifying limitations and possible areas for improvements to support the process. Therefore, the questions are divided into the following categories: the art-making process, the artwork (product), mastery of the system, and general experience. The rest of the discussion is guided by observations made during the individual workshop. By utilizing both quantitative and qualitative methods, the study aims to provide a detailed understanding of how participants interpret and engage with AI-generated visual art.

Results

A total of 20 participants were involved in the study, with 10 participants in each condition. The text-to-image condition consisted of 10 participants aged 18-29, including 6 females and 4 males. The sketch-to-image condition also had 10 participants aged 18-29, comprising 8 females, 1 male, and 1 non-binary individual.

Affect Grid

Condition/Time points(n=10)	T1		T2		T3	
	Valence	Arousal	Valence	Arousal	Valence	Arousal
text-to-image	5.70(1.57)	5.55(1.21)	5.40(1.57)	6.50(1.37)	7.20(1.00)	6.95(1.54)
sketch-to-image	7.10(1.29)	5.45(1.77)	7.25(1.78)	6.20(1.03)	6.20(2.46)	5.50(1.67)

Table 1: Mean valence and arousal scores (and standard deviations) at times t1, t2, and t3 for each condition.

Text-to-Image Condition:

A repeated measures ANOVA test for valence scores at T1, T2, and T3 for the text-to-image condition shows $p=0.001$.

Post hoc comparisons reveal significantly different valence scores from t2 to t3 ($p=0.002$) and from t1 to t3 ($p=0.006$). This suggests that participants' valence improved after completing co-creating tasks with text-to-image generative AI.

Another repeated measures ANOVA test for arousal scores at T1, T2, and T3 shows $p=0.042$. Post hoc tests reveal that the significant increase in arousal occurred between T1 and T3 ($p=0.044$). This suggests that participants experienced a notable rise in arousal after finishing both tasks.

Sketch-to-Image Condition

Same tests were conducted for valence and arousal scores for the sketch-to-image condition. No significant change in valence was observed for the sketch-to-image condition ($p=0.413$). Similarly, there was no significant change in arousal scores ($p=0.393$), although changes in mean score demonstrated slight improvement in valence(+0.15) and arousal(+0.75) between t1 and t2. The drop in valence from t2 to t3 may have resulted from a drastic decrease in valence score from two participants (from 9 to 2 and 8 to 3). Overall, the sketch-to-image condition did not experience notable variations in either valence or arousal throughout the experiment.

Comparison

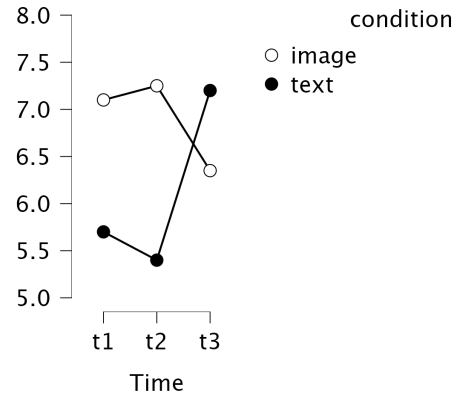


Figure 4: Valence Score Over Time

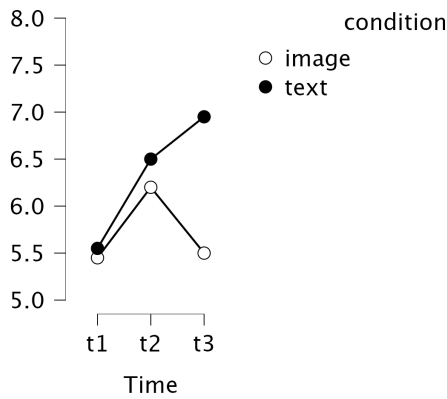


Figure 5: Arousal Score Over Time

An independent samples t-test reveals a significant difference in valence prior to the tasks ($p=0.021$), with the sketch-to-image condition exhibiting higher pre-experiment valence compared to the text-to-image condition. However, there is no significant difference in arousal levels before the tasks.

The findings support the hypothesis that participants in the text-to-image condition experienced therapeutic effects, as evidenced by significant improvements in both valence and arousal scores. This suggests that participants experienced improved mood and an increase in activation after engaging in text-based co-creation with AI.

With higher initial valence, participants in the sketch-to-image condition did not exhibit significant changes in either valence or arousal scores after the tasks. Therefore, the hypothesis is not supported for the sketch-to-image condition. Potential reasons for these findings will be explored in the Discussion section.

Flow Short Scale

Condition/ Time points (n=10)	T2			T3		
	Total	Absorb	Fluency	Total	Absorb	Fluency
text-to-image	46.00(6.68)	17.30(2.91)	27.80(6.02)	50.10(6.28)	19.10(2.56)	31.10(4.89)
sketch-to-image	48.10(13.07)	18.70(4.60)	29.20(8.98)	50.60(14.28)	20.20(5.41)	30.40(8.98)

Table 2: Mean flow score and subscores (and standard deviations) at T2 and T3 for each condition. Flow score changes were positive in both conditions from T1 to T2; subscores for fluency and absorption for both conditions also improved.

Text-to-Image Condition

In the text-to-image condition, the individual items with the highest mean scores were: "My thoughts/activities run freely and fluidly," "I have no difficulty concentrating," and "I'm totally absorbed in what I'm doing." These items suggest that participants in this condition experienced particular elements of flow state, characterized by ease of thinking, high concentration, and absorption in the task.

Paired sample t-test suggests a significant increase ($p=0.049$) in flow total score after task 2 for the text-to-image condition. This corresponds with the increase in valence and arousal scores from time t2 to t3, as greater enjoyment may be related to deeper immersion and engagement.

Sketch-to-Image Condition

For the sketch-to-image condition, the items with the highest mean scores were: "I have no difficulty concentrating," and "I'm totally absorbed in what I'm doing," confirming the high concentration deep absorption experienced by participants. However, the overall flow score did not significantly change between time t2 and t3. This suggests that while participants felt concentrated and absorbed, these feelings did not intensify over time as they did in the text-to-image condition. In addition, one participant in the sketch-to-image condition submitted much lower scores for both Affect Grid and FSS, which could have impacted the overall results. This highlights the variability in individual experiences within the sketch-to-image condition.

Comparison

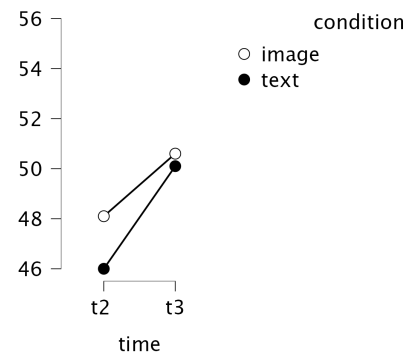


Figure 6: Flow Total Score Over Time

When comparing the two conditions, both showed an improvement in flow scores from time t2 to t3, with mean flow scores falling into the medium range. This indicates a trend towards enhanced flow states after finishing task 2, and both conditions can induce a moderate level of flow. However, the repeated measures ANOVA revealed no significant difference in flow scores between the two conditions, suggesting that neither condition was superior in enhancing flow.

In summary, the hypothesis that co-creating with AI leads to improved mood and flow state is more strongly supported for the text-to-image condition, with clearer indications of improved affect. The sketch-to-image condition shows some positive trends but lacks strong statistical support.

Interviews

After completing the experiment, every participant took part in a short interview. The interview questions are divided into the following categories: artwork (product), art-making, mastery of the system, and general experience. The rest of the discussion is guided by observations made during the individual workshop.

Category	Example Questions
Art	Does the outcome match with your expectation of the event/emotion you wanted to portray? Do you think it helps you express your feelings?
Art-making	How did you come up with the prompt? Did drawing or putting into words help or hinder extracting your experience and emotion? Does viewing and modifying the images help you discover anything related to yourself?
Mastery of the system	Did you encounter any difficulties in any of the steps? What were the steps that disrupted the process, and what were the steps that you enjoyed?
General Experience	What improvements or changes do you expect for a better experience?

A thematic analysis, including familiarization, initial codes, searching and reviewing themes, and finalization (Braun and Clarke, 2006), was conducted on the interview data after transcribing the audio recordings. The results for each category are grouped into themes based on the RQs.

Main Themes	Sub-Themes/Description
Art (Artifact): 1. Control	1.1 Challenges in control and requirement for skill and experience 1.2 Mixed satisfaction 1.3 Realization of limitations
2. Inconsistent Image Quality	2.1 Aesthetic perception 2.2 Inconsistent accuracy

3. Varied Attitudes	3.1 Open-minded and explorative 3.2 Goal-oriented tolerance 3.3 Cycle of expectation and motivation, learning curve
---------------------	---

Art-making:

RQ1: Were enjoyment, self-expression and reflection achieved?

1. Expression through AI	1.1 Incomplete but acceptable portrayals 1.2 Creative style exploration
2. Reflections Achieved Through Visualization and Prompt Editing	2.1 Concrete visualization facilitated reflection and clarification 2.2 Experimentation with prompts both assisted and hindered expression
3. New Perspectives, Highlighted Emotions, Limited Insight	3.1 Good representation but limited insight 3.2 New perspectives through variation and reimagination
RQ2: How can co-creating process support enjoyment, self-expression, and reflection?	
4. Dialogue and Iteration with AI	4.1 Exploration and Curiosity 4.2 Mixed feelings for communication with AI
5. Separate Roles of Input, Output, and Tasks	5.1 Role of sketching 5.2 Role of text prompts 5.3 Role of AI outputs 5.4 Realizations and Discoveries through Variations and Prompt Editing 5.5 Goal and Task Comparison

Mastery of the System and General Experience:

Tool (Features and Biases)

Interesting strategies

Suggested Improvements

Table 3: Interview Thematic Analysis

Discussion

RQ1: Does co-creating visual art with AI demonstrate therapeutic potential through improved affect, facilitating flow experience, and achieving self-expression and reflection?

The quantitative results suggest that the text-to-image condition demonstrated therapeutic potential in terms of improved valence and arousal scores. In contrast, the results from the sketch-to-image condition were less consistent. While both conditions achieved a moderate flow state, high concentration and absorption in the task stood out as the most prominent aspects of their experiences. This is further confirmed by the interviews, where participants reported a smooth overall experience with no significant disruptions or frustrations. Many participants found the process enjoyable and fun, especially when they were open to letting AI be creative. While the AI-generated images didn't always perfectly match their imagination, they captured key elements that effectively reflected their intended experiences or emotions. Participants noted that expression and self-reflection were achieved to some extent, where balancing control and freedom in the creative process was crucial.

There are several possible reasons for the insignificant improvement in valence for the sketch-to-image condition.

Participant-specific factors: Two participants showed a drastic drop in valence between t2 and t3. One of them reported feeling "depressed" due to the task of recalling a nightmare, which likely influenced her emotional state negatively and impacted the overall results. Her observation was echoed by other participants who shared that imagining scenes related to personal stressors can bring back stressful memories, making the task less enjoyable. Another participant, who engages with art professionally and has strong skepticism towards AI, reported very low flow state scores. Removing this outlier, the sketch-to-image condition shows a significantly higher absorption score at t2 as compared to the text condition ($p=0.033$, independent t-test), indicating that professional experience and pre-existing attitudes towards AI might affect engagement levels.

Technical issues: Dreamstudio demonstrated poorer performance in interpreting sketches. The resulting images tended to be either exact replicas of the original sketches or entirely new and unrelated, leading to higher uncertainty and a perceived lack of control among participants. Additionally, the need to take a photo of their drawing and upload it required a shift between interfaces, which may have disrupted the flow and contributed to a less seamless and enjoyable experience.

Expression through AI:

While participants are open to letting the AI decide on details, they also expressed a necessity for essential elements to be accurately represented—the overall vision, feeling, and atmosphere needed to capture similar emotions. Elements that were not initially considered, such as specific objects or symbols, emerged during the creative process; these unexpected elements are often embraced if they resonate with the overall atmosphere or personal vision.

Participants experimented with various styles primarily for fun, aesthetic appeal, and emotional expression. Some participants aim for maintaining consistency in storytelling across images and achieving specific emotions, while others focus on the enjoyment and stimulation derived from testing different styles, rather than on emotional alignment. Therefore, self-expression is partially achieved through co-creating images with AI, though in a more exploratory manner, where experimentation and fun are emphasized over emotional expression during the process.

New Perspectives, Highlighted Emotions, Limited Insight:

Participants mainly focused on achieving a specific image, the process hence provides a good representation of their mental image but revealed limited insight into their experiences. The short timeframe of the experiment also limited significant emotional change or meaning-making. However, participants commented on how the process led to new perspectives, through identifying missing elements from image variations, highlighting crucial emotional adjectives in the prompt, or reimagining a negative scenario.

Reflections Achieved Through Visualization and Prompt Editing:

Concrete visualization facilitated clarification and reflection by enabling participants to compare and evaluate visible representations of their internal states. One participant noted that having tangible images to depict past or future thoughts offered a clear visual contrast between different timeframes. The ability to view different variations also allowed users to refine their mental images, making them more specific and concrete over time. While the resulting images might not always align perfectly with their initial intentions, the process of generating variations, editing prompts, and evaluating results encouraged reflection on previously unexamined aspects and helped users identify what they genuinely wanted to see in the final image. Overall, the process of creating and reviewing the generated images contributed to a better understanding

and clarification of their thoughts; participants appreciated it as an interesting alternative to introspection alone.

Experimentation with prompts can be both rewarding and challenging. Firstly, the process of editing and refining prompts helps participants in identifying and correcting missing or unwanted elements, which enhances the precision of the mental image and makes the desired output clearer. Secondly, participants were able to recognize repeated use of specific terms in their prompt, such as "lonely," revealing deeper, previously unnoticed aspects of their thoughts and feelings. Lastly, the iterative process of prompt editing often fosters excitement, motivation, and realization of goals. However, adapting to text prompt requirements may hinder one’s spontaneous expression.

Table 4: Examples of Text Prompt Development

Development of One Prompt	
cave with darkness cave with darkness. messy. empty cave with darkness. messy. with one light empty cave with darkness. messy. with one star	Person that is discovering their dark and light inner world Woman that has a little colorful spot near her heart but the rest of the body is black A single and little burst of color against a black stormy sky, in a body of woman
Prompt Development for Reimagination	
apocalypse dark nightmare with dragon breathing fire, city skyscrapers, burning, smoke, destruction, dark night, flying dragons in the background	A homemade short film shooting in the netherlands gone wrong
apocalypse cute sparkly pastel nightmare dragon breathing fire, cute village farm meadow, sheep, flower field, sunny, cloud, flying dragons in the background	A cinema screen with a film from a canal in the Netherlands A cinema screen with a film of a couple of friends biking in a cloudy dutch canal containing slight orange and purple tones A outdoor cinema screen showing the successful airing of a dutch short film, containing a real image of a cloudy canal



Figure 7: Examples of Image Variations, Symbols, and Reimaginings

To sum up, observations from the interviews indicate that creating images with AI fosters enjoyment through experimentation with different aesthetic options, facilitates self-expression through personally meaningful visualizations, and encourages reflection by clarifying and refining mental images across variations. Both quantitative and qualitative results support the therapeutic potential of generating visual art through AI, particularly in the text-to-image condition. Collaborating with AI for visual art-making offers a promising avenue for enhancing mood, self-expression, and reflection, provided that the beneficial factors are preserved and the undesirable qualities avoided, which will be discussed in the next section.

RQ2: How can co-creating with image generation AI support enjoyment, self-expression, and reflection?

Dialogue and Iteration with AI

Participants reported that it was easy to enter a state of exploration and fun experimentation. They enjoy adding elements, exploring different styles, and varying their approaches to see what unique outcomes emerge. Facilitating this curiosity-driven experimentation can enhance the overall creative experience.

Participants showed mixed feelings about their “dialogue” with the AI tool. The AI presents its interpretation of the user’s sketches and/or prompts, and the user responds by making selections and refinements, making the process more engaging and interesting. While the experimentation process gives surprising and enjoyable results, the lack of control and frequent misinterpretations by the AI can be stressful. Some participants even described it as “fighting against the AI” or intentionally challenging it to achieve the desired images.

Most participants intuitively adopted an incremental approach in their creative process. After evaluating generated images, they gradually added or removed elements to achieve clearer and more satisfying outcomes. The iterative nature of working with AI involves varied prompt writing strategies. Participants may focus on main elements, emotional adjectives, or creative guesses about what might work. Most see the process as a cooperative effort, where adapting to the AI’s language helps better articulate their vision. One participant likened her interaction with the AI to collaborating with another person, noting that trying to communicate with the AI can potentially lead to frustrations, but reacting to its generated images contributed to concretizing her ideas. These interactions help clarify and concretize what participants have in mind, leading to a better understanding of their thoughts and expectations. This “mutual interpretation” (Liu et al. 2024), where participants and AI engage in a

back-and-forth exchange, is valued by some as the most enjoyable aspect of the process.

However, not all iterations lead to improvement; some may regress, requiring rewrites and retries, making the unpredictability both exciting and frustrating. Therefore, the process should aim for encouraging a sense of progression, allowing participants to recognize that their modifications lead to increasingly nice and relevant outcomes.

Separate Roles of Input, Output, and Tasks

Role of sketching

Participants find that beginning with a manual drawing allows them to maintain a certain level of control and deliberate gestures that AI alone cannot provide. Some participants prefer the conscious process provided by manual drawing, using AI as a complement to the drawing skills they may lack. Sketching thus served as a starting point for AI to add details or offer new interpretations that might be more cumbersome to achieve manually.

Sketching also played a pivotal role in clarifying ideas and allowing participants to focus on essential elements. It assisted participants in understanding specific images and feelings they had, simplifying complex mental concepts into core components. Participants viewed this translation of mental images into clear, simple elements in their sketches as an effective way to generate text prompts for AI. Exploring the multimodal approach can hence not only improve output quality but also provide a more engaging experience by helping participants achieve control and focus.

Role of text prompts

Text prompts played a complicated, sometimes conflicted role in the co-creative process. Participants use text prompts to both express themselves and communicate with AI. Describing their ideas in words is seen as straightforward and necessary for refining the generated images; however, refining these descriptions to craft effective prompts can hinder spontaneity and emotional expression. Some participants found translating complex and abstract ideas into text prompts challenging in the first place. Additionally, the process of making these prompts fit for AI involves trial and error, which can be logical and less spontaneous, potentially hindering the fluidity in emotional expression. However, for some participants, the process of refining prompts a fun and creative exercise that leads to new insights and improved results.

Sometimes, writing is preferred for it is a more familiar process than sketching, and it captures details better.

Overall, crafting a prompt for the AI tool to produce a good image requires strategy and skill. Guidance and direction is needed to retain its constructive side in clarifying and concretizing thoughts, rather than restricting expressions.

Role of AI outputs

AI significantly lowers the barriers to artistic creation. Specifically, it is effective in quickly presenting a close approximation of participants' envisioned concepts. This capability helps in quickly generating a rough visual that can be refined further, saving considerable time and effort compared to traditional methods. This makes the creative process less daunting and more inviting, particularly for those who may not have much art-making experience. Despite its advantages, AI often falls short in capturing the intensity of emotions that users wish to convey. This limitation leads some users to prefer manual drawing, as it ensures a more precise representation of their vision and feelings. The personal touch and control inherent in hand-drawing are sometimes necessary to achieve the desired emotional depth and authenticity in the artwork. This prompts questions on the specific contexts or purposes for which AI or traditional art-making approaches are more preferable.

Realizations and Discoveries through Variations and Prompt Editing

The process of generating and viewing variations involves exploring both the technical capabilities of the AI and discovering new inspirations. The decision-making involved in selecting and adjusting variations helps participants gain a deeper understanding of their own thoughts and feelings. The selection process is guided by a pre-existing mental image, following which participants make changes and adjustments to improve inaccurate, repetitive, or unsatisfactory AI results to reach a final output that aligns more closely with their vision. This feedback loop between visualization, evaluation, and modification encourages continuous exploration and refinement, often leading to realization of missing elements or new inspirations.

Several participants observed that viewing visual representations makes thoughts and ideas that were natural and instinctive to oneself more explicit. However, one needs to be cautious that participants may prefer to accept visually appealing results rather than striving for a closer depiction of their ideas. This is echoed by observations from Du et al. (2024) where users strive for sophisticated visuals over pursuing self-expression. Therefore, it is crucial to balance the pursuit of aesthetic satisfaction with the goal of achieving self-expression and understanding.

Goal and Task Comparison

The co-creative process is influenced by personal preferences and nature of the tasks, with clear objectives enhancing engagement and satisfaction. Having a specific task can generate anticipation, excitement, and sense of achievement, particularly when the outcome aligns with the user's vision. Some participants indicated that they are more likely to use AI with a clear purpose in mind, rather than aimlessly experimenting; they prefer tasks that have personal relevance over trying randomly shuffled prompts. However, when users aim for a very specific and concrete image tied to personal experiences, AI-generated results often fall short, leading to frustration due to misinterpretations. Moreover, tasks tied to personal experiences can deeply affect users' moods, as recalling and drawing unpleasant scenarios may evoke distress.

For open-ended, exploratory tasks, participants have a more relaxed attitude towards the final output, and are more accepting of surprising results. However, the freedom offered by the non-directive approach may be confusing for participants who don't consider themselves very creative, as they found it challenging to come up with prompts. Lack of direction at the beginning can thus be slightly frustrating as users need to search for their own focus. For them, tasks in Part 2 are more enjoyable because of a clearer connection to the participants' personal experiences.

The task approach—whether specific or open-ended, positive or negative—significantly influences the user's experience. Open-ended tasks are generally more enjoyable and less stressful, fostering discovery, creativity and engagement. Specific tasks, while potentially bringing frustrations, can lead to a greater sense of fulfillment and more meaningful outcomes when they align well with users' personal experiences and goals. For mood improvement, the distraction strategy in therapeutic art-making may similarly work better for human-AI co-creation; for self-expression and reflections on their personal experiences, which serve different goals in the directive approach in art therapy, it will entail different strategies and considerations, depending on the goal.

Implications

Another hypothesis was that the sketch-to-image would experience more pronounced therapeutic effects for it retained some aspects of traditional art-making activities, such as embodiment and non-verbal expression, though it was partially rejected by the quantitative results. Apart from the aforementioned factors that may have skewed the results, this discrepancy raises questions about whether co-creating with AI operates on a shared set of therapeutic dynamics with different priorities. What characterizes

art-making experience with AI, and where do creative enjoyment and reflections happen?

The first point to highlight, as many participants also came to realize, is that AI is unlikely to ever produce an image perfectly aligned with their vision. The goal of co-creating images with AI is therefore not to produce a flawless representation but to engage in interactions that facilitate creative exploration and clarification of their ideas. This aligns with the approach of “emphasi[zing] artmaking as a healing process, ... [instead] of using art as merely a product for item analysis to depict mental states” (Du et al., 2024). In a way, the inaccuracy and lack of spontaneous, personal touch in homogenized AI-generated images rejects being analyzed for its formal qualities.

In this study, the user is both the creator collaborating with AI and the viewer who constantly evaluates and responds to AI-generated artifacts. The discoveries they made may seem less significant or noticeable due to the absence of an art therapist who usually acts as a facilitator of “self-expression and reflection through the process of art making and the resulting artwork” (de Witte et al., 2021). With this limitation in mind, the interactions and qualities that demonstrated the potential of being “the third hand,” which facilitates the creative process without being imposing (Pease et al., 2022), can be summarized as recommendations for incorporating human-AI co-creative processes into therapeutic interventions.

What stood out the most from the co-creative process was enjoyable experimentation and gradual concretization. Depending on the goal, different aspects of the two key characteristics can be emphasized.

1. Balance freedom and control to enhance mood and flow state

The state of intrinsically motivated exploration and experimentation is one of the most prominent features of the co-creative process with AI. One participant linked the inaccurate or unexpected aspects of generated images to an incomplete space on canvas that encourages creativity and personalization, allowing them to gradually build and refine their ideas. Findings by Kwon et al. (2021) confirmed that imperfections and serendipity significantly enhance satisfaction with AI collaborative creations. Similarly, the study by Liu et al. (2024) identified AI's unconditional acceptance of imaginative concepts and its ability to offer endless variations as empowerment for clients in a family therapy context. Therefore, the surprising possibilities offered by AI generations can be leveraged as a stimulus for achieving enjoyment and flow state.

However, too much freedom can be disorienting. In this case, control can be supported through involving sufficient user input and providing a clearer sense of progress and accomplishment. Du et al. (2024) cautioned against full automation and emphasized the importance of ensuring the users' participation, thus, allowing users to see more visible consequences of their input and actions, and providing a clearer structure or goals may be helpful. Viewing this process in the light of casual creators, text-to-or sketch-to-image AI generations offer a vast possibility space that needs more restrictions to ensure user's confident exploration (Compton and Mateas, 2015). Therefore, we can consider incorporating features that limit possible actions and reinforce a positive feedback loop in the tools, such as direct manipulation and instant feedback.

2. Encourage dialogue within and across to achieve reflection.

In the co-creative process, concretization is achieved through variations of prompts and visualizations. Participants transform their internal states into visible representations and gradually carve out a more specific image by refining text prompts. The back-and-forth between AI and user not only induces deep immersion and absorption, but also engages the user in internal and external dialogues through which adjustment and decision-making in language and image choice can uncover overlooked feelings and insights.

However, some level of external support is often necessary to help formulate expressions, identify patterns, make observations, ask questions, and assist in discoveries. Du et al. (2024) evaluated AI based on the ETC model and emphasized its power on the cognitive side, as differentiated from intuitive art-making with traditional materials. Therefore, reminder or guidance is needed to ensure its support for clarifying and concretizing thoughts, rather than restricting expressions. Similarly, Wan et al. (2024) recommend offering scaffolding to support both the creation and recall of elusive experiences. Besides the process, images serve as a great starting point to elicit verbal responses, helping to extract personal symbols and metaphors. In this case, images are used as artifacts to encourage or moderate conversation. Again, introducing another viewer, whether a trained therapist or not, can potentially facilitate expression and communication.

3. Venting or Distraction? Different Task Designs

As demonstrated in the discussion above, participants showed different reactions and preferences to the different task designs. It is then helpful to revisit the different traditional art-making approaches and goals employed in therapeutic contexts to discuss which tasks may serve better under specific contexts.

For short-term mood improvement, co-creating with AI demonstrates great potential for it improves valence, arousal and flow, especially when participants view it as a fun, exploratory task to create neutral or positive images, potentially as distraction. While tasks of personal relevance can be more engaging and give a greater sense of achievement, caution needs to be taken to reduce frustration caused by AI misunderstanding and consolidate appreciation of new, positive perspectives. Just as Gruber and Oepen (2018) observed in their narrative synthesis of emotion regulation strategies and effects in art-making, art therapy interventions are most effective in alleviating negative mood states when they entail a structured task leading to a specific outcome or goal state. Incorporating image generative AI in therapy contexts is promising for extracting symbols, discovering new aspects, and encouraging reflections, but more research is needed to address specific implications.

4. Multimodality in both input and output

Participants appreciated sketching's role in bridging the gap between conceptualization and digital representation. This multimodal interaction not only facilitates a more effective realization of their artistic vision, but also enriches the creative process as participants engage in more diverse ways to brainstorm ideas, experiment with different strategies, and clarify their goals. Some participants suggested having more art-making materials to start with may improve their experience, as they have a greater sense of control and involvement for the input images.

Liu et al.'s (2024) study on generative AI and family therapy suggested the materialization of generative AI results to within and beyond therapy context. This is echoed by participants who expressed strong interest in printing the results for decoration, visual diary, or even future counseling sessions. Therefore, expanding the types of input and output involved in the co-creative process holds great potential for improving user experience and discovering new catalysts for therapeutic effects.

5. Addressing issues inherent to general AI tools

Participants reported several technical issues, such as unresponsive negative prompt, image strength, and random inappropriate content flags. In addition, AI systems may exhibit inherent racial and gender biases or cultural misrepresentations. This can significantly affect the therapeutic value experienced by the participants. Hence, addressing these technical and ethical issues in AI tools is crucial for ensuring a positive and effective user experience.

Limitations and Future Work

Generalizability: Apart from the limited sample size, participants of this study consists mostly of university students who are likely to be more familiar with technology and open to explore new digital tools. However, technical barriers can exist for specific groups. This also means there is room for discussing how human-AI co-creation can be adapted to different mental health challenges and target groups.

Additionally, Dreamstudio, the AI tool used in the study was not specifically designed for therapeutic purposes and lacked tailored design considerations. Critical factors outlined in Pease et al.'s A Roadmap for TCC (2022), such as "Matching the client to the medium" "Collaborate with mental health professionals" and "Design software which is underpinned by work in art therapy" are therefore not addressed.

Missing Therapist Perspective: The study lacked the involvement of art therapists, which could have provided valuable insights into how the generated artifacts and the art-making process are interpreted in a therapeutic context. Beyond the art (artifacts), art-making (process assisted by AI), client (user), the therapist and client-therapist relationships can be a crucial therapeutic factor (de Witte et al., 2021). The absence of a therapist's perspective means that the potential impact of the AI-generated artifacts on therapist-client communication was not explored.

Potential Novelty Effect: Though this was mitigated by the time given for trying out the tool before the experiment, the study may have been influenced by a novelty effect, where participants' reactions and experiences were shaped by the newness of the AI tool rather than its intrinsic therapeutic value.

Ethical issues: The AI tool's unpredictable outputs could lead to unintended or distressing content, posing risks in a therapeutic setting. Additionally, relying on AI for creative processes might undermine user autonomy, reducing their sense of ownership and control in the artmaking and therapeutic experience. These ethical concerns necessitate careful consideration in designing and deploying AI tools for mental health applications.

This study separated the discussion of therapeutic potential into therapeutic art-making and art-making used in art therapy, which fitted with Recommendation 3 from A Roadmap for Therapeutic Computational Creativity, "Distinguish therapeutic from therapy" (Pease et al., 2022). In addition, the combination of quantitative and qualitative measures to assess impacts on mood and flow state, and to define possible therapeutic factors, offers a plausible approach for evaluating effectiveness of therapeutic

human-AI co-creation. Though it again would have benefited from incorporating the therapist's perspective, as the Roadmap suggested, merging evaluation criteria from both disciplines is valuable, and necessary, for developing suitable and practical metrics for TCC.

Drawing insights from short-term mood repair, the ETC model, and therapeutic factors for CAT, the therapeutic potential of co-creating visual art with AI is thus characterized as intrinsically motivated exploration and experimentation, and dynamic dialogue in the art-making process, where the art, the art-maker, and the AI tool contributes to synergistic iteration and feedback. As a result, human-AI co-creation demonstrates therapeutic potential in mood regulation, enhancing flow state, and facilitating expression and reflection. To build upon these findings, explore and enhance the application of AI in therapeutic contexts, future work can focus on the following areas:

Developing and Evaluating Strategies and Tasks: Investigate how different strategies and tasks can be optimized for various settings. For instance, consider the different roles that sketching, writing, and other materials or strategies for creating and manipulating the input and output may play in facilitating creativity, enjoyment, and expression.

Specific Groups and Contexts: Examine the applicability of AI tools across specific groups, including different health conditions, age groups, and contextual settings. Some tasks or strategies may be particularly interesting depending on requirements from specific groups. For example, distraction and pleasant imagery may be more effective depending on preferences and traits of the target group. Tailoring approaches to these variables will ensure more relevant and effective applications.

Individual Differences: Individual differences observed in the sketch-to-image group highlighted the need for personalized therapeutic interventions. Understanding how personal experiences and attitudes impact engagement and effectiveness of the co-creative process can thus inform how and what co-creation activities could cater to different preferences and needs. Future research can explore how individual preferences and traits, such as previous experience with art-making, familiarity with AI, and personal attitudes, influence the effectiveness and outcomes of human-AI co-creation in therapeutic contexts.

Tool Improvement: Develop tools informed by user feedback and research findings. Enhance the tool by simplifying the interface for both verbal and visual inputs, improving control over specific image areas, and addressing other technical issues.

Longitudinal Studies: Investigate whether users continue to engage with the tools regularly and assess the sustained effects on their creative and therapeutic outcomes.

Conclusion

Creation as a profession or labor shifts the emphasis to quality and efficiency; similarly, creativity categorized as a personality trait, a compartmentalized function, or a god-like power can discourage many who may say, "I'm not an artist, I can't draw well." Findings of this study redirect the focus away from speed and quality and illustrate that sometimes a spontaneous mark on paper can mean more than a polished image. Therefore, this study is an investigation of possible uses of generative AI tools that do not prioritize efficiency and allow therapeutic creativity to happen. Exploring the interconnectedness between casual creators, image generative AI, and therapeutic practices, this study highlights enjoyable experimentation and gradual concretization as the therapeutic value of human-AI co-creation. Balancing freedom and control in exploration optimizes mood and flow, while encouraging the triadic dialogue between art, user, and AI promotes expression and reflection.

Acknowledgements

This project would not be possible without the generous support from many mentors and friends. I thank Rob for reassuring me with his story of finding a good question after three years; his patient guidance was important in helping me find a meaningful question for this project. I am also grateful to Tessa for her continuous support throughout this journey. Thank Yiming and Ray for helping with experiments and being my best roommates in the Hague, the film, food, and library trips we had together truly brightened my days. Thank Xiao for working with me on pain coping skills gamification and supporting me through various projects, which led to the initial idea of this study. Thank Iris for helping with the pilot test and lending her brilliant friends.

A big thank-you to every participant in the study who generously shared their experiences and told me they had enjoyed the session. To my family, thank you for always being supportive even when we are separated by hours and borders. Special thanks to my mother, Fang, who taught me the importance of care and exploration. Thank everyone who kindly offered me time for a conversation throughout this journey. Every discussion, documented or not, has enriched this project with surprises, challenges, reflections, and fun.

References

- Angheluta, A.-M., & Lee, B. K. (2011). Art therapy for chronic pain: Applications and future directions. *Canadian Journal of Counselling and Psychotherapy*, 45(2). <https://cjc-rcc.ucalgary.ca/article/view/59289>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Compton, K., & Mateas, M. (2015). Casual creators. In *Proceedings of the Sixth International Conference on Computational Creativity* (June 2015). Retrieved from <https://axon.cs.byu.edu/ICCC2015proceedings/10.2Compton.pdf>
- Dalebroux, A., Goldstein, T. R., & Winner, E. (2008). Short-term mood repair through art-making: Positive emotion is more effective than venting. *Motivation and Emotion*, 32, 288–295. <https://doi.org/10.1007/s11031-008-9105-1>
- de Witte, M., Orkibi, H., Zarate, R., Karkou, V., Sajjani, N., Malhotra, B., Ho, R. T. H., Kaimal, G., Baker, F. A., & Koch, S. C. (2021). From therapeutic factors to mechanisms of change in the creative arts therapies: A scoping review. *Frontiers in Psychology*, 12, 678397. <https://doi.org/10.3389/fpsyg.2021.678397>
- Drake, J. E., Coleman, K., & Winner, E. (2011). Short-term mood repair through art: Effects of medium and strategy. *Art Therapy*, 28(1), 26–30. <https://doi.org/10.1080/07421656.2011.557032>
- Drake, J. E., & Winner, E. (2012). Confronting sadness through art-making: Distraction is more beneficial than venting. *Psychology of Aesthetics, Creativity, and the Arts*, 6(3), 255–261. <https://doi.org/10.1037/a0026909>
- Du, X., An, P., Leung, J., Li, A., Chapman, L. E., & Zhao, J. (2024). DeepThInk: Designing and probing human-AI co-creation in digital art therapy. *International Journal of Human-Computer Studies*, 181, 103139. <https://doi.org/10.1016/j.ijhcs.2023.103139>
- Futterman Collier, A., & Wayment, H. A. (2021). Enhancing and explaining art-making for mood-repair: The benefits of positive growth-oriented instructions and quiet ego contemplation. *Psychology of Aesthetics, Creativity, and the Arts*, 15(2), 363–376. <https://doi.org/10.1037/aca0000286>
- Gruber, H., & Oepen, R. (2018). Emotion regulation strategies and effects in art-making: A narrative synthesis. *The Arts in Psychotherapy*, 59, 65–74. <https://doi.org/10.1016/j.aip.2017.12.006>
- Han, A., & Cai, Z. (2023). Design implications of generative AI systems for visual storytelling for young learners. *IDC '23: Proceedings of the 22nd Annual ACM Interaction Design and Children Conference*, 470–474. <https://doi.org/10.1145/3585088.3593867>
- Hinz, L. D. (2019). *Expressive therapies continuum: A framework for using art in therapy*. Routledge.
- Hogan, S. (2016). *Art therapy theories: A critical introduction*. Routledge.
- Jeon, Y., Jin, S., Shih, P. C., & Han, K. (2021). FashionQ: An AI-driven creativity support tool for facilitating ideation in fashion design. *CHI '21: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, Article No. 576, 1–18. <https://doi.org/10.1145/3411764.3445093>
- Kaimal, G., Mensinger, J. L., Drass, J. M., & Dieterich-Hartwell, R. M. (2017). Art therapist-facilitated open studio versus coloring: Differences in outcomes of affect, stress, creative agency, and self-efficacy. *Canadian Art Therapy Association Journal*, 30(2), 56–68. <https://doi.org/10.1080/08322473.2017.1375827>
- Kim, K. S., Loring, S., & Kwekkeboom, K. (2018). Use of art-making intervention for pain and quality of life among cancer patients: A systematic review. *Journal of Holistic Nursing*, 36(4), 341–353. <https://doi.org/10.1177/0898010117726633>
- Kwon, O., Bae, S., Lee, D., Lee, K., Choi, R., Hwang, K., Kwon, S., & Kim, T. (2024). An empirical study on the psychological improvement effects and satisfaction of Korean traditional painting generative AI. In *2024 International Conference on Artificial Intelligence in*

- Information and Communication (ICAIC)* (pp. 135-140). <https://doi.org/10.1109/ICAIC60209.2024.10463404>
- Lee, Y. K., Park, Y. H., & Hahn, S. (2023). A portrait of emotion: Empowering self-expression through AI-generated art. *Annual Meeting of the Cognitive Science Society*, April 26, 2023. <https://doi.org/10.48550/arXiv.2304.13324>
- Literat, I. (2013). "A pencil for your thoughts": Participatory drawing as a visual research method with children and youth. *International Journal of Qualitative Methods*, 12(1), 84-98. <https://doi.org/10.1177/160940691301200143>
- Liu, D., Zhou, H., & An, P. (2024). "When he feels cold, he goes to the seahorse"—Blending generative AI into multimaterial storymaking for family expressive arts therapy. CHI '24: Proceedings of the CHI Conference on Human Factors in Computing Systems, Article No.: 118, 1–21. <https://doi.org/10.1145/3613904.3642852>
- Pease, A., Ackerman, M., Pease, N., & McFadden, B. (2022). A roadmap for therapeutic computational creativity. Paper presented at 13th International Conference on Computational Creativity, Bolzano, Italy.
- Phillips, J., Ogden, J., & Copland, C. (2015). Using drawings of pain-related images to understand the experience of chronic pain: A qualitative study. *British Journal of Occupational Therapy*, 78(7), 404-411. <https://doi.org/10.1177/0308022614562791>
- Rankanen, M. (2016). Clients' experiences of the impacts of an experiential art therapy group. *The Arts in Psychotherapy*, 50, 101-110. <https://doi.org/10.1016/j.aip.2016.06.002>
- Rheinberg, F., Vollmeyer, R., & Engeser, S. (2003). *Flow Short Scale* [Database record]. APA PsycTests. <https://doi.org/10.1037/t47787-000>
- Rosal, M. L., & Gussak, D. E. (2015). Cognitive-behavioral art therapy revisited. In M. L. Rosal & D. E. Gussak (Eds.), *The Wiley handbook of art therapy* (pp. 68-76). John Wiley & Sons, Ltd.
- Russell, J. A., Weiss, A., & Mendelsohn, G. A. (1989). Affect grid: A single-item scale of pleasure and arousal. *Journal of Personality and Social Psychology*, 57, 493–502.
- Schwarzer, R., & Jerusalem, M. (1995). *General Self-Efficacy Scale (GSE)* [Database record]. APA PsycTests. <https://doi.org/10.1037/t00393-000>
- Singh, B. (2011). The therapeutic effects of art making in patients with cancer. *The Arts in Psychotherapy*, 38(3), 160-163. <https://doi.org/10.1016/j.aip.2011.04.001>
- Stickley, T., Wright, N., & Slade, M. (2018). The art of recovery: Outcomes from participatory arts activities for people using mental health services. *Journal of Mental Health*, 27(4), 367–373. <https://doi.org/10.1080/09638237.2018.1437609>
- Suprpto, N., Sunarti, T., Suliyanah, Wulandari, D., Hidayaatullaah, H. N., Adam, A. S., & Mubarak, H. (2020). A systematic review of Photovoice as participatory action research strategies. *International Journal of Evaluation and Research in Education*, 9(3), 675–683.
- Sun, J., Yang, J., Zhou, G., Jin, Y., & Gong, J. (2024). Understanding human-AI collaboration in music therapy through co-design with therapists. CHI '24: Proceedings of the CHI Conference on Human Factors in Computing Systems, Article No. 704, 1–21. <https://doi.org/10.1145/3613904.3642764>
- Van Lith, T., Fenner, P., & Schofield, M. (2010). The lived experience of art making as a companion to the mental health recovery process. *Disability and Rehabilitation*, 33(8), 652–660. <https://doi.org/10.3109/09638288.2010.505998>
- Wan, Q., Feng, X., Bei, Y., Gao, Z., & Lu, Z. (2024). Metamorpheus: Interactive, affective, and creative dream narration through metaphorical visual storytelling. CHI '24: Proceedings of the CHI Conference on Human Factors in Computing Systems, Article No. 166, 1–16. <https://doi.org/10.1145/3613904.3642410>
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive

and negative affect: the PANAS scales. *Journal of personality and social psychology*, 54(6), 1063.

Yoo, D., Kim, D. Y. J., & Lopes, E. (2023). Digital art therapy with Gen AI: Mind palette. In *2023 11th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW)*.
<https://daki7711.github.io/files/yoo2023digital.pdf>