

Master Thesis

Chatbot-Assisted Guidance in Digital Drawing:

Evaluating AI chatbot features on enhancing user's creativity and experience.

Author:

Leila Matayeva

Faculty 3 - Mathematics/Computer Science, Department of Digital Media

University of Bremen

June 7th, 2024

Examiners:

Dr.-Ing. Nina Wenig

Prof. Peter von Maydell

Faculty 3 - Mathematics/Computer Science

Faculty of Art and Design

University of Bremen

Hochschule für Künste Bremen

Bibliothekstr. 5, 28359 Bremen, Germany

Am Speicher XI 8, 28217 Bremen

Acknowledgement

I would like to express my sincere gratitude to the University of Bremen for providing an excellent educational environment during my Master's studies.

I am deeply grateful to my first supervisor, Dr.-Ing. Nina Wenig, for her continuous support, guidance, and encouragement, as well as to my second supervisor Prof. Peter von Maydell for his valuable advice and insightful critiques throughout the thesis process.

I extend my gratitude to the participants of my user studies for their time and input, which greatly contributed to my research.

Lastly, I am very thankful to my family and friends for their support, patience and belief in me.

Abstract

This master's thesis presents a comprehensive study on the influence of chatbot features on users' creativity and user experience in a drawing web application. Three different versions of the application were produced, each with a different set of AI enhanced features of the chatbot to support and enhance users' creative process while drawing. Two user studies, using questionnaires and ratings of creativity of produced artworks, performed statistical analysis in determining the efficiency of different chatbot features in inspiring, engaging, and supporting users' creativity.

The findings hinted at a very remarkable difference between different versions of the application in terms of the perceptions and creative productions that would contribute to the understanding of the effects of AI-powered chatbots in digital drawing environments towards fostering creativity.

TABLE OF CONTENT

1.	<i>Introduction</i>	1
1.2	AI-enhanced chatbots and their role in creative fields.....	2
1.3	Problem statement and motivation.....	3
1.4	Research questions.....	4
2.	<i>Background</i>	6
2.1	Chatbots' role in enhancing user experience and creativity	6
2.2	Previous studies on AI-supported creativity in digital art	7
2.3	Existing Tools	10
3.	<i>Methodology</i>	14
3.1	Development of the Drawing Web Application	14
3.2	Development of the three chatbot versions.....	15
3.2.1	Active Chatbot	16
3.2.2	Recognizing Chatbot.....	18
3.2.3	Drawing Chatbot.....	19
3.3	Database integration for collecting data	21
3.4	Design of the User Studies.....	21
3.4.1	First User Study	22
3.4.2	Second user study	25
3.5	Statistical analysis methods	28
4.	<i>Results</i>	30
4.1	First user study	30
4.1.1	Demographics	30
4.1.2	CSI results	30
4.1.3	CUQ results.....	32
4.1.4	User interaction	33
4.1.5	Comparative Analysis	34
4.1.6	Correlations.....	36
4.1.7	Qualitative analysis	37
4.1.8	Additional Analysis	39
4.2	Second user study	41
4.2.1	Demographics	41
4.2.2	CAT results	41
4.2.3	Comparative analysis	42
5.	<i>Discussion</i>	45
5.1	Assessing the Influence of Chatbot Features on Supporting Users' Creativity	45

5.2	Evaluating User Experience.....	47
5.3	Limitations of the Study.....	48
6.	<i>Conclusion</i>	50
6.1	Key Findings	50
6.2	Future Design of AI-Enhanced Chatbots in Digital Drawing.....	52
7.	<i>References</i>	54
8.	<i>Appendices</i>	57
8.1	Appendix 1. Consent form	57
8.2	Appendix 2. Data protection form	58
8.3	Appendix 3. Demographics questionnaire for the first user study.....	59
8.4	Appendix 4. First User Study Questionnaire for each version	62
8.5	Appendix 5. End questionnaire for the first user study	73
8.6	Appendix 6. Demographics questionnaire for the second user study	78
8.7	Appendix 7. Consensual assessment technique questionnaire	80
8.8	Appendix 8. Demographics answers for the first study	93
8.9	Appendix 9. Answers to the ‘Active’ chatbot version questionnaire (CSI, CUQ, additional questions)	94
8.10	Appendix 10. Answers to the ‘Recognizing’ chatbot version questionnaire (CSI, CUQ, additional questions)	97
8.11	Appendix 11. Answers to the ‘Drawing’ chatbot version questionnaire (CSI, CUQ, additional questions)	100
8.12	Appendix 12. Answers to the second user study	103

1. INTRODUCTION

In recent years, the intersection of artificial intelligence (AI) and creative activities has attracted an increasing interest. Among various creative domains, digital drawing stands out as one of fields where AI can play a notable role. One crucial development in this area is the creation of chatbots, with AI-enhanced chatbots standing out as powerful tools for enhancing creative processes. These chatbots offer real-time assistance, feedback, and inspiration to users [1]. Despite these benefits of using AI chatbots in creative fields, there is a lack of research on how different chatbot features specifically impact users' creativity and user experience. This master's thesis aims to investigate this topic and contribute to the academic research in this area

1.1 DIGITAL DRAWING AND CREATIVITY

Creativity is a fundamental human ability that promotes innovation, problem solving, and personal fulfilment. Studies have proven that people are able to achieve better cognitive flexibility, nurture their emotional well-being, and experience higher levels of life satisfaction through creative activities [2], [3]. One of the most effective ways to develop and enhance creativity is through the creation of art. Artistic processes, such as drawing, provide a safe space for experimentation and risk-taking, which are important parts of the creative process [4].

Drawing, whether digital or traditional, is an inherently creative process that involves several cognitive functions simultaneously, such as problem solving, planning, and spatial reasoning. Some research reports claim that it enhances creativity by activating the visual and motor areas of the brain. Drawing also encourages divergent thinking, which leads artists to generate different ideas and possible solutions [5]. Moreover, the act of drawing helps with emotional expression and regulation, providing a therapeutic outlet for stress and anxiety. This emotional engagement further enhances creative thinking by allowing artists to freely explore their inner thoughts and feelings [6].

In recent years, digital drawing has gained popularity as a new way of artistic expression, often replacing traditional drawing methods [3]. The exploration of digital art started in the 1960s as a

new form of media art [3]. Digital drawing increased its popularity starting in the 1980s and 1990s, as more advanced and sophisticated software programs such as Adobe Illustrator and Photoshop were introduced, giving artists powerful tools for creating digital artwork. Moreover, the introduction of internet platforms for artists revolutionised digital drawing by enabling artists to share their work worldwide and collaborate online with others [7].

While both digital and traditional drawing share the core principles of art, they differ in several aspects. Traditional drawing uses physical media, while digital drawing uses digital devices and software. Digital drawing allows easy corrections and modifications, whereas traditional drawing requires more effort to change [8]. As a result, digital drawing started to attract high interest, offering variety and accessibility to a wide audience.

Studies have shown that the integration of various digital tools, which are widely used in digital drawing, can further enhance the creative process by providing artists with a unique opportunity to visualise and realise their ideas [9]. Nowadays, professional artists as well as beginners are able to use digital tools to unleash their creativity and express themselves without the limitations of traditional media.

1.2 AI-ENHANCED CHATBOTS AND THEIR ROLE IN CREATIVE FIELDS

Artificial Intelligence (AI) is developing at a rapid pace, and nowadays tools with AI capabilities are becoming increasingly popular in various fields, including creative ones such as writing, music composition and art creation. AI has transformed these fields by providing innovative solutions that improve productivity, creativity and engagement.

A number of applications and websites demonstrate how these AI-based tools are integrated into the creative domain. For example, Artbreeder¹ is an online application that allows users to blend different artistic images using AI-driven prompts. In writing, for example, Grammarly² and Jasper³ both use artificial intelligence to help writers check grammar and make style suggestions for the content being written. Similarly, in music composition there is AIVA⁴ (Artificial Intelligence Virtual Artist) application, where musicians can create new tracks either by sampling existing

1. Artbreeder. (n.d.). Artbreeder. Retrieved from <https://www.artbreeder.com/>

2. Grammarly. (n.d.). Grammarly: Free Online Writing Assistant. Retrieved from <https://www.grammarly.com/>

3. Jasper (n.d.). Jasper. Retrieved from <https://www.jasper.ai/>

4. AIVA (n.d.). AIVA, the AI Music Generation Assistant. Retrieved from <https://www.aiva.ai/>

work or even allowing users to give their own input on what kind of song they want, it then uses algorithms based on user input.

AI tools are also used in some digital art creation platforms, supporting user engagement and promoting creativity [10]. For example, Adobe Fresco⁵ integrates such technologies into the interface so that artists may receive suggested variations of their drawings, as well as auto-fill sketches or colour palettes for their current work.

Amongst the variations of these tools there are chatbots powered by AI, that are often integrated into digital platforms. AI-powered chatbots can serve as virtual assistants, providing inspiration, feedback and technical guidance to artists and creators. Using advanced AI technologies, these chatbots are able to simulate human conversation and interact with users in a natural and intuitive way. These chatbots can understand and respond to user input, make recommendations and provide real-time assistance. The potential of AI-powered chatbots to revolutionise the creative industries is becoming increasingly clear [1].

1.3 PROBLEM STATEMENT AND MOTIVATION

Despite the potential benefits of AI-enhanced chatbots in creative activities, there is limited empirical evidence on how different chatbot features specifically impact users' creativity and user experience. To design instruments that would really support drawing process it is important to recognize these impacts. Recent research highlights the importance of interaction and feedback in creative work, suggesting that AI tools that focus on supporting human creativity rather than replacing it can be extremely beneficial [11].

Currently, most AI applications in the creative domain focus primarily on creating final products, rather than assisting in the creation phase [12]. Therefore, filling this gap by integrating AI-enhanced chatbots into digital drawing platforms could provide continuous support during the art creation process [10].

Hence, there is a need to investigate how digital tools such as chatbots can be designed considering the crucial role of creativity for innovation, problem solving and personal fulfilment. Therefore,

5. Adobe. (n.d.). Adobe Fresco: Draw and paint your way. Retrieved from <https://www.adobe.com/products/fresco.html>

the motivation for this study stems from the integration between AI technologies, chatbots and drawing platform, coupled with the need to investigate their impact on user creativity and user experience. This research aims to understand how different aspects related to chatbots influence the artistic process, thus contributing to the development of more efficient digital tools for artists. The significance of this study lies in its potential to enhance the user experience in digital drawing applications, contributing to the broader field of human-computer interaction and creativity support systems.

1.4 RESEARCH QUESTIONS

The main research questions of this master's thesis are:

I. Assess the Influence of Chatbot Features on Supporting Users' Creativity:

RQ1: How do various chatbot functionalities support users' creativity while using a digital drawing application?

H1: Different chatbot features will support users' creativity in varying degrees.

H0: There will be no difference in the support for users' creativity provided by different chatbot features.

II. Evaluate User Experience

RQ2: What is the impact of chatbots with different features on overall user experience while using a digital drawing application?

H2: Chatbots with different features will result in different user experiences levels.

H0: There will be no difference in user experience and satisfaction across chatbots with different features.

To design the features of chatbots for further user tests, a thorough review of existing literature and current tools was conducted. This process identified the key strengths of important AI chatbot features as well as existing gaps. Chatbot features differ in their technical functionalities and interaction styles. The insights gained will inform the design of more effective AI-enhanced chatbots, ultimately contributing to the advancement of creativity support systems in digital drawing applications.

Based on these research questions, this master's thesis explores the influence of different chatbot features on users' creativity and user experience in a digital drawing web application. It begins with a short introduction to the topic and literature review, examining existing research papers and current tools in the field. The methodology section gives details to the selection of technologies for developing the web application and the implementation of chatbots with varied features. This is followed by the design and execution of two user studies. Finally, there is a comprehensive analysis of the collected data, including a discussion of the results and their significance.

2. BACKGROUND

Creativity Support Tools (CSTs) have long been researched with the aim of enhancing human creativity through technology. Shneiderman (2007) emphasised the importance of CSTs in accelerating discovery and innovation by providing tools for the creative process in various domains such as art, music, writing and others [10]. Artificial intelligence has made many recent advances, particularly in relation to its use in creative activities. AI-enabled chatbots are one such innovation that has attracted much attention due to its potential impact on user creativity and user experience in various artistic domains. This part outlines previous studies on chatbots and AI tools used in CSTs to support creativity and enhance user experience, addressing both achievements and gaps in the literature.

2.1 CHATBOTS' ROLE IN ENHANCING USER EXPERIENCE AND CREATIVITY

Chatbots, often referred to as conversational agents, are transforming the way industries handle knowledge and interactions. The way industries manage knowledge bases is being transformed by chatbots, or 'conversational agents'. Driven by advances in artificial intelligence such as Natural Language Processing (NLP), Deep Learning (DL), Machine Learning (ML), etc., chatbots are continuously evolving for better functionality and usability [13]. Due to recent advances such as OpenAI's GPT series, which belongs to the group of deep learning models, chatbots can now handle more complex tasks with context awareness [14].

Chatbots are applied across various fields, including e-commerce, healthcare, education, and entertainment [14], [15]. They are proving to be more efficient than static web pages when it comes to effectively engaging users [16]. There are two main ways to implement chatbots: rule-based systems and AI-based models. Rule-based chatbots rely on predefined rules as well as pattern matching to generate responses; they are easier to design but cannot handle complex queries sufficiently [17]. In contrast, AI-based models are interactive systems that learn from interactions and improve their performance in time. They usually consist of a user interface, a dialogue manager, an (NLP) engine and a database to store information and responses [18].

According to Haase and Hanel (2023), chatbots are becoming integral to how we work, communicate, and create. They enhance user experience (UX) by providing immediate, personalised, and efficient interactions [19]. Chatbots have the ability to personalise the recommendations and solutions offered to users by ensuring that they are always relevant to the needs of the individual. Furthermore, the availability of chatbots on a continuous basis increases the overall satisfaction as well as the convenience [18].

Chatbots encourage creative problem-solving and critical thinking by offering diverse perspectives and instant feedback. Professionals in the creative industries use them for brainstorming sessions, content creation or generating ideas for new concepts and approaches [20]. As previously noted by Haase and Hanel (2023), a comparison of human-generated ideas with those generated by generative AI chatbots, including versions of ChatGPT, showed no qualitative difference in creativity [19], while Urban et al. (2023) found that working on complex creative tasks alongside chatbots can improve individuals' self-efficacy for creative performance [21]. By interacting with chatbots, users can explore a wider range of ideas and approaches, inspiring new thoughts and solutions.

Although chatbot technologies have evolved considerably, several challenges persist in chatbot development. Capturing the contextual and emotional nuances of human conversations remains difficult, affecting the ability to engage users effectively [17]. There is growing interest in developing chatbots that can dynamically adapt to users input for more personalised interactions [17]. Integrating multimodal inputs, for example a visual or auditory cues, can create more immersive and interactive experiences, enhancing the chatbot's ability to understand and respond to user needs [19].

2.2 PREVIOUS STUDIES ON AI-SUPPORTED CREATIVITY IN DIGITAL ART

AI tools have been widely used in creative fields. Early systems, such as Harold Cohen's AARON, and more contemporary examples, such as Creative Adversarial Networks (CAN), have shown how AI tools are evolving in the art field. CAN develops artworks by acquiring styles through

training and intentionally deviating from them to explore the limits of what is considered creative [12]. These advances have facilitated the development of AI as a collaborator, rather than just a production tool.

The idea of co-creativity, a situation in which humans and AI can interact, has been widely explored. The way artists perceive their artistic process has been transformed by the integration of AI into creative workflows. Indeed, according to Davis et al. (2021), co-creative systems can provide real-time feedback, suggestions and modifications. Such interaction helps to overcome creative blocks and promote new ideas, leading to an overall increase in creativity. For example, platforms such as Drawing Apprentice engage users in collaborative drawing, supporting co-creation where both humans and computers contribute equally to drawings. In these cases, participatory sense-making enables AI contributions that are dynamically integrated into the user's creative workflow [22]. Another example can be systems like StoryDrawer [23] and CharacterMeet [24], which have shown how AI can facilitate creative storytelling and character development. These systems use AI to generate sketches and suggest narrative elements, helping users overcome creative blocks and explore new ideas. The interactive nature of these tools allows for a collaborative experience, where AI augments human creativity by providing novel suggestions and inspirations.

Another example are AI-enhanced educational drawing tools that have been shown to be effective in promoting creativity. The AI-based drawing system studied by Xu et al. notably increased user engagement, with a corresponding increase in creativity. Similarly, the integration of artificial intelligence into education has been shown to promote higher levels of creativity and engagement through personalised support and instant feedback, such as the AI-based student drawing system developed by Hartati & Manullang [25]. These tools take advantage of AI tools that provide instant help throughout the creation process, making it more enjoyable for learners. However, their use in professional artistic practices should be further explored [23], [26].

While various studies have explored the role of AI in co-creative systems, many of them specifically highlight the potential of chatbots and AI agents to assist artists. Study by Lynda Clark (2021) explores the potential use of AI as a creative assistant, suggesting ideas and improving the

creative process [27]. Schmitt et al. (2021) propose the use of AI-enhanced chatbots as virtual assistants that enhance creative processes by simulating human-like conversations and providing contextual guidance [28].

Several studies point to the importance of user's engagement in co-creative systems. For instance, research by Kantosalo et al. (2021) on participatory sense-making in abstract drawing with a co-creative cognitive agent emphasises the need for systems that support meaningful interaction and engagement. The study found that users value the ability to interact dynamically with AI systems, which can lead to more satisfying and productive creative experiences [29]. Bickmore et al. (2005) found that relational agents that initiate interactions and maintain a conversation can improve user adherence and satisfaction [30]. Another empirical study by Caldarini et al. (2021) provides a comprehensive review of AI-driven chatbots, highlighting their application in various creative fields. The authors identify key factors that influence the effectiveness of chatbots in supporting creativity, such as user engagement and the quality of AI-generated outputs [13].

Another important aspect of AI-enhanced chatbots is their ability to provide user-chatbot collaboration and feedback. Ibarrola et al. presented one such successful example, the CICADA system. CICADA is a context-sensitive drawing agent designed to assist professional designers in the early stages of conceptual exploration. The system enriches creativity through dynamic adaptation to user input and a collaborative design process [31]. A number of studies have investigated the role of AI-driven feedback mechanisms in enhancing creativity. The work of Hwang and Won (2021) explores how social facilitation through AI chatbots can enhance creative idea generation. Their findings suggest that chatbots can support divergent thinking and provide valuable feedback that stimulates creative processes [32].

Despite some advancements of implementing new technologies in digital drawing systems, there is still not enough research made in this area. Therefore, several challenges remain in the effective integration of chatbots in creative processes. Although there are different levels of interactivity provided by existing systems, more intuitive and engaging interfaces are needed. For example, Sun et al.'s (2021) study of the ORIBA system demonstrated the importance of dynamic interaction. However, achieving continuous meaningful engagement in the creation of visual art is

problematic [33]. Zhang et al. also showed that AI often loses relevance with context over extended interactions. Another research on StoryDrawer also highlights the need for AI tools to adapt dynamically to user contributions to keep the context relevance [23].

One more critical challenge in human-computer co-creation is balancing the initiative of the user and the AI chatbot. Mixed-initiative creative interfaces, such as Sentient Sketchbook [34], imply the need for systems capable of changing their level of intervention in response to user input. However, designing such systems remains complex, and further work is needed to develop intuitive and responsive AI collaborators [34]. Moreover, there is a lack of standardised evaluation metrics for assessing the impact of AI on creativity. The reviewed studies therefore suggest the need for reliable methods to measure creativity outcomes and user satisfaction. For example, Haase & Hanel's (2023) research on AI-driven creativity highlights a reliable evaluation framework as an important aspect [19].

The body of research reviewed provides valuable insights into the capabilities and limitations of AI-enhanced chatbots in creative processes. Based on previous research, we can emphasise the importance of aspects which highly affect the user's experience, such as adaptability to user inputs, collaboration, engagement and feedback. While noteworthy progress has been made, particularly in terms of interactivity, feedback mechanisms, and user engagement, there are critical areas that need further exploration. Many studies have primarily focused on the technical capabilities of AI and its applications without thoroughly exploring how specific features of AI-enhanced chatbots affect user creativity and experience. For example, while systems like CoCo Sketch [35] have shown how AI can support creative exploration through real-time interaction, there is limited empirical evidence on how different chatbot features, such as feedback styles and personalization, influence the creative process [35]. Therefore, this master's thesis offers a digital drawing application which has AI-enhanced chatbots with different features to study the nuanced aspects to have the best result in digital drawing environments in terms of creativity and user experience.

2.3 EXISTING TOOLS

Despite a large number of AI tools and APIs available for creative activities, there is a remarkable pattern: more attention is given to the final product than to the actual creative process. Many tools

can turn sketches into amazing paintings, but they often focus on the end result rather than the journey of creation. For instance, DeepDream by Google⁶ stands out as a pioneering neural network model designed to generate captivating and often surreal artistic images by iteratively enhancing patterns found in existing images. Similarly, DoodleGAN⁷ uses advanced generative algorithms to transform simple doodles into complex drawings by adding details and textures that captivate the eye. FastStyleTransfer⁸, meanwhile, encourages users to experiment with different artistic styles by seamlessly transferring the visual characteristics of well-known works of art to their own digital creations. Another popular API is Pix2Pix⁹, available in both TensorFlow and PyTorch, enables the generation of realistic and visually appealing artwork from simple outlines by learning the mapping between input sketches and detailed images. Moreover, Neural-Style-Transfer¹⁰, implemented in PyTorch, empowers users to infuse their images with the stylistic characteristics of renowned artists, resulting in unique and visually attractive compositions. Although these tools attract a high interest, users may find themselves limited to predefined styles or constrained by the capabilities of the algorithms, leaving little room for personal exploration and artistic growth.

Conversely, there is a limited number of tools and APIs that centre on the creative process itself. They include tools such as Magic Sketchpad, Color Pop, AutoDraw and Draw.AI. Magic Sketchpad¹¹, for example, provides intelligent guidance for completing drawings by recognising the user's strokes and making suggestions for refinement or improvement based on computer vision technology. Color Pop¹² is able to create colouring books through prompts that allow its users to explore with different colour palettes and composition options. Another popular application is AutoDraw¹³, which speeds up the creation process with its intuitive suggestions that enrich the visual narrative of one's artwork with various objects generated based on the user's sketches. While the idea behind these applications is highly interesting, some of them fail to be successfully implemented. For example, Draw.AI¹⁴ works like an interactive mentor that guides you step-by-step through drawing lessons, offering valuable techniques that are illustrated during each lesson.

6. Google. DeepDream. Retrieved from <https://github.com/google/deepdream>
7. DoodleGAN. Retrieved from <https://github.com/jpark73/DoodleGAN>
8. TensorFlow. Fast Style Transfer. Retrieved from https://www.tensorflow.org/tutorials/generative/style_transfer
9. TensorFlow. Pix2Pix. Retrieved from <https://www.tensorflow.org/tutorials/generative/pix2pix?hl=ru>
10. PyTorch. Neural Style Transfer. Retrieved from https://pytorch.org/tutorials/advanced/neural_style_tutorial.html
11. Magic Sketchpad. Retrieved from <https://magicsketchpad.com/>
12. Color Pop. Retrieved from <https://colorpop.com/>
13. AutoDraw. Retrieved from <https://www.autodraw.com/>
14. Draw.AI. Retrieved from <https://draw.ai/>

However, testing of the application revealed issues with drawing functionalities. The example of Draw.AI application usage is shown on image 1:



Image 1 – Example of Draw.AI application

There are also various APIs available. One example is DoodleNet¹⁵, which is based on TensorFlow and focuses on doodle recognition, a foundation for applications that use doodles as the main interaction and creativity. In addition, Sketch-RNN¹⁶ uses recurrent neural networks to produce diverse and realistic-looking drawings, helping artists to create images that are always true to themselves through sequential stroke generation.

Similarly, Google's QuickDraw¹⁷ seeks to engage users' interest in spontaneity by inviting them to sketch freehand while the AI attempts to identify their drawing in real time, with the aim of enhancing user interaction with drawing interfaces that encourage creativity through play. The example of using QuickDraw application is shown on image 3:

15. DoodleNet. Retrieved from <https://github.com/googlecreativelab/quickdraw-dataset>
16. TensorFlow. SketchRNN. Retrieved from https://magenta.tensorflow.org/assets/sketch_rnn_demo/index.html
17. Google QuickDraw. Retrieved from <https://quickdraw.withgoogle.com/>

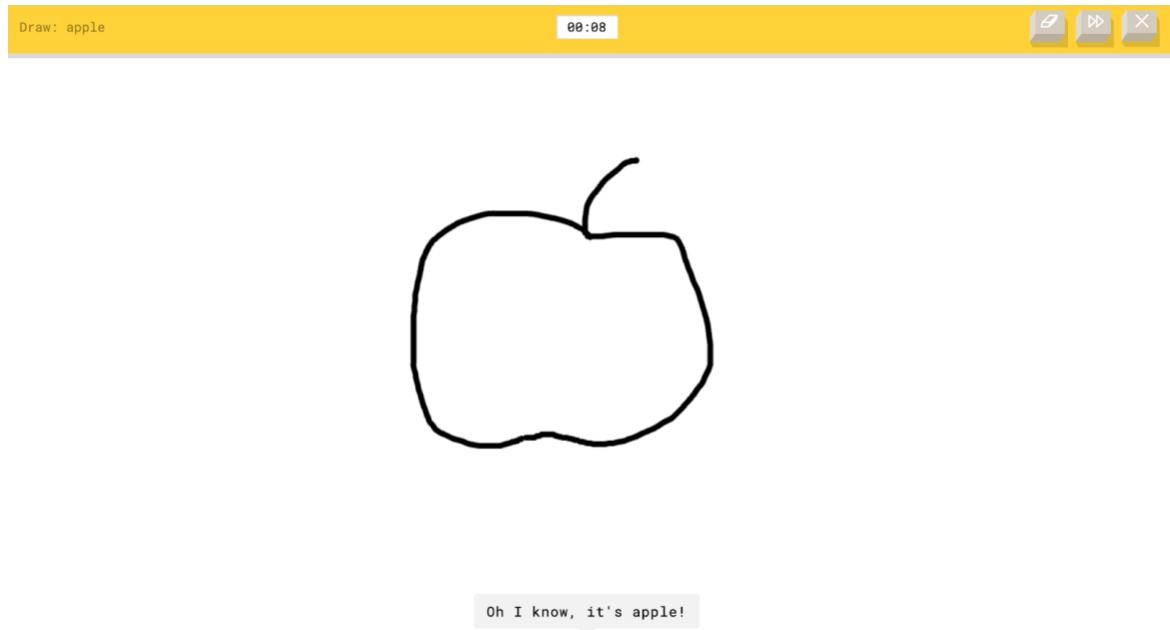


Image 2 – Example of QuickDraw! Application

Given the focus of my master thesis on enhancing user creativity and experience through digital drawing, I have chosen to incorporate appropriate APIs such as SketchRNN and QuickDraw into my work.

3. METHODOLOGY

Based on extensive literature and existing tools research, it was decided to create a drawing web application with basic drawing functionalities. Within the application, three different versions of AI chatbots were integrated, each tailored to support and enhance the users' creative drawing experience in unique ways. Based on their functionalities, they would further be mentioned as 'Active', 'Recognizing' and 'Drawing' chatbots. An SQLite database tracked user interactions within the app. After technical implementation, to evaluate the chatbots' effectiveness, two user studies were conducted: the first involved participants using all three chatbot versions to create drawings, followed by questionnaires assessing their experiences; the second involved rating the creativity of the produced artworks. The details of all steps are described further in this chapter.

3.1 DEVELOPMENT OF THE DRAWING WEB APPLICATION

The decision to develop a web-based drawing application was driven by the need for accessibility and ease of use. The web application was developed using the Flask framework, along with JavaScript, HTML, and CSS. Flask, a lightweight web framework for Python, was chosen for its simplicity and flexibility, which facilitated rapid development and easy integration of various features. JavaScript was used to handle the dynamic aspects of the application, such as interactive drawing functionalities, while HTML and CSS were employed to structure and style the web pages respectively. This combination of technologies ensured that the application was both functional and user-friendly. For the whole development PyCharm IDE, developed by JetBrains, which is perfectly suited for web frameworks and databases.

The application was designed with a simple and intuitive interface to minimise learning curves and encourage creativity. It consists of a drawing canvas, toolbar and a chatbot. The toolbar, positioned prominently on the interface, offers several basic drawing tools:

- Simple drawing brush
- Colour and size chooser, allowing users to select their desired colour and size of the stroke.
- Eraser for correcting mistakes.
- 'Clear all' button to start over with a blank canvas.

This straightforward and simple design was intentional, since the focus of the study is on chatbots functionalities, rather than on the drawing application itself. By providing essential tools without overwhelming users with options, the application helps users focus on their creative process rather than the tool itself. The finished result of the developed web application can be seen on the image 3 below:



Image 3 - Drawing web application

The development of the drawing web application, combined with a simple user-friendly interface, created an accessible and engaging platform for studying the influence of chatbot features on users' creativity.

3.2 DEVELOPMENT OF THE THREE CHATBOT VERSIONS

In developing the chatbots for this study, each version was designed with distinct features that aim to engage users in creative process, inspire exploration of new ideas and provoke them to collaborate with the chatbot - aspects needed for enhancing user creativity and experience. For the development of these features, it was chosen to use open-source AI APIs. When initiating a specific version of the application, the chatbot would send an initial message explaining its

capabilities to the user. An example of initial message from the ‘Recognizing’ chatbot is on the image 4 below:

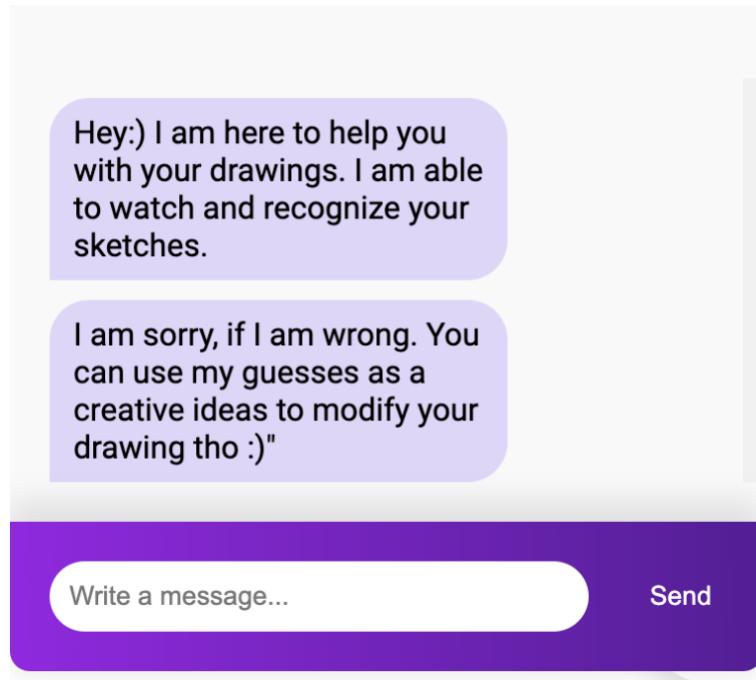


Image 4 - Example of the initial message

The three versions—‘Active’, ‘Recognizing’, and ‘Drawing’—each offered unique functionalities to enhance users’ creative experiences.

3.2.1 ACTIVE CHATBOT

The main feature of the ‘Active’ chatbot was its proactive and initiative behaviour. Using the ChatGPT API (GPT-3.5 Turbo)¹⁸, this chatbot actively engaged with users by sending messages independently, asking about their progress, offering advice, and providing ideas for what to draw. The primary goal was to simulate the presence of a real person constantly interacting with the user as well as inspire the user to engage in conversation with the chatbot, exploring new ideas and tips for drawing. If the user was inactive for more than 10 seconds, the chatbot would check in to ensure everything was okay.

The ChatGPT API was specifically tuned to provide short, friendly, and humorous responses, acting as a drawing teacher by offering precise step-by-step advice. Additionally, the chatbot had

18. OpenAI. ChatGPT API Retrieved from <https://openai.com/index/introducing-chatgpt-and-whisper-apis/>

a set of predefined messages containing general drawing advice collected from several drawing books such as “Drawing for the Absolute and Utter Beginner” by Garcia C. [36] and “Drawing: A Complete Guide” by Civardi G. [37], which it would send every few minutes to maintain engagement and inspiration. Example of conversation with ‘Active’ chatbot version can be seen on image 5:

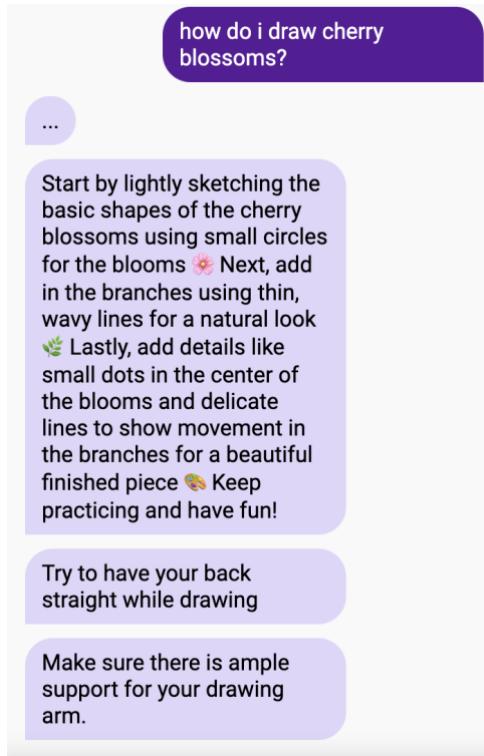


Image 5 - Example of conversation with Active Chatbot

The ‘Active’ Chatbot was designed to be a proactive and engaging presence within the drawing web application. This chatbot is characterised by its initiative behaviour, frequently interacting with users to enhance their creative process. The motivation behind creating such a chatbot stems from previously described research indicating that proactive and initiative AI systems can notably boost user engagement and creativity. The key aspects of ‘Active’ Chatbot’s features are:

- Enhanced Engagement: by providing frequent interactions and initiating conversations. The constant presence and interaction can prevent users from feeling stuck or unmotivated, encouraging continuous creative exploration

- Continues support: by sending messages with drawing tips and ideas that simulate an art teacher, that checks on a user's progress and gives general advice.

3.2.2 CHATBOT

The ‘Recognizing’ Chatbot’s unique feature was its ability to identify what the user was drawing. After every fifth stroke, the chatbot attempted to guess the user’s drawing using the QuickDraw API by Google, which includes 345 recognition categories. This feature aimed to make users feel acknowledged and ‘seen’, and inspire them to experiment with different drawings, either by validating the chatbot’s guesses or by drawing something new based on incorrect guesses. Example of using ‘Recognizing’ chatbot version is shown on the image 6 below:

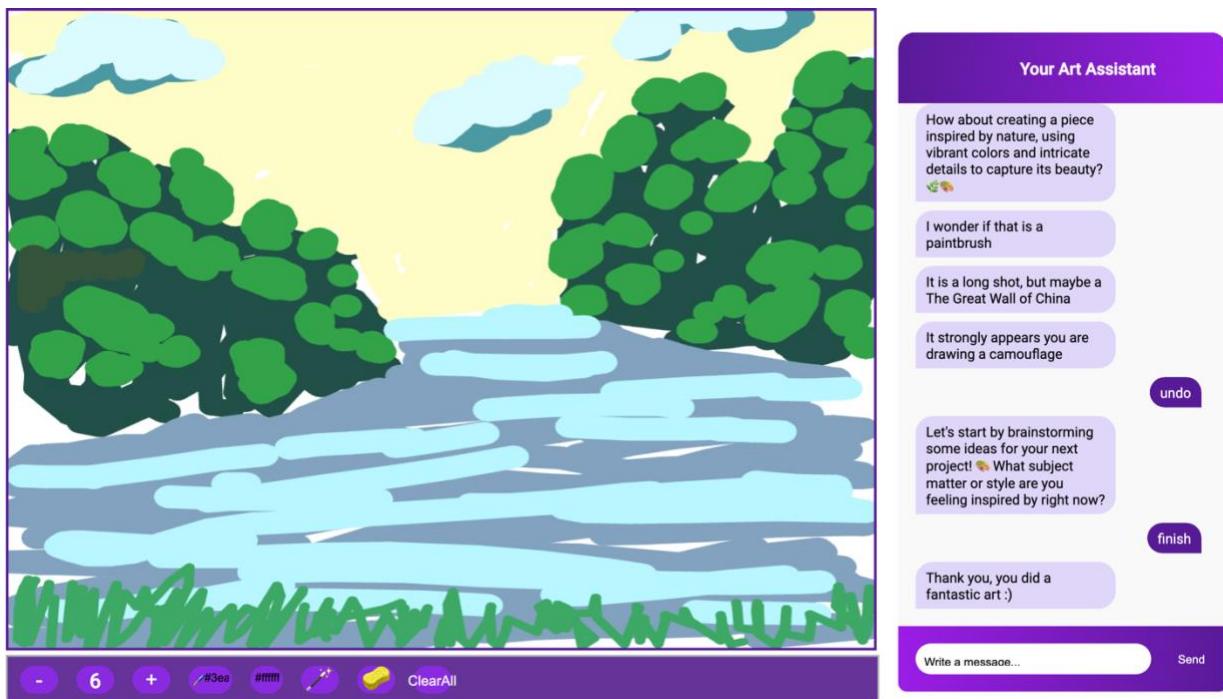


Image 6 - Example of using ‘Recognizing’ Chatbot

This chatbot provided feedback based on its confidence level in the recognition:

- 90%-100%: "I am absolutely certain you are drawing a..."
- 80%-90%: "I am pretty confident you are trying to draw a..."
- 70%-80%: "It strongly appears you are drawing a..."
- 60%-70%: "I am fairly sure you are sketching a..."
- 50%-60%: "Looks like you might be drawing a..."

- 40%-50%: "I am guessing you are trying to draw a..."
- 30%-40%: "Could it be that you are drawing a..."
- 20%-30%: "I wonder if that is a..."
- 10%-20%: "It is a long shot, but maybe a..."
- 0%-10%: "I am not sure, but it could possibly be a..."

This chatbot also used the ChatGPT API to respond to user messages but did not send unsolicited messages apart from its drawing guesses. It continued guessing until it detected a change in the user's drawing subject, at which point it would pause its guessing. The key aspects of 'Recognizing' Chatbot's features are:

- Enhancing Engagement: By recognizing and responding to users' drawings, the chatbot maintains a dynamic and interactive environment, and make users feel 'seen', which can increase user engagement and motivation
- Encouraging Exploration: The chatbot's guesses encourages users to experiment with different drawing subjects, increasing a desire to discover and explore drawing new objects.

3.2.3 DRAWING CHATBOT

The 'Drawing' chatbot's primary function was to directly participate in the user's drawing process by adding lines to the artwork. Initially, it would suggest drawing a specific object. Once the user began, the chatbot would, after several strokes, contribute by adding complementary lines. Users could instruct the chatbot to draw something different, delete its last strokes or to stop altogether.

The idea behind this chatbot was to inspire users to approach their drawings in novel ways and to foster a sense of collaboration with the chatbot. This version utilised the SketchRNN API, which recognizes the user's lines and continues them. While it also employed the ChatGPT API for communication, it did not send unsolicited messages, instead it focused on the collaborative drawing process. The example of how 'Drawing' chatbot version was used can be seen on image 7:

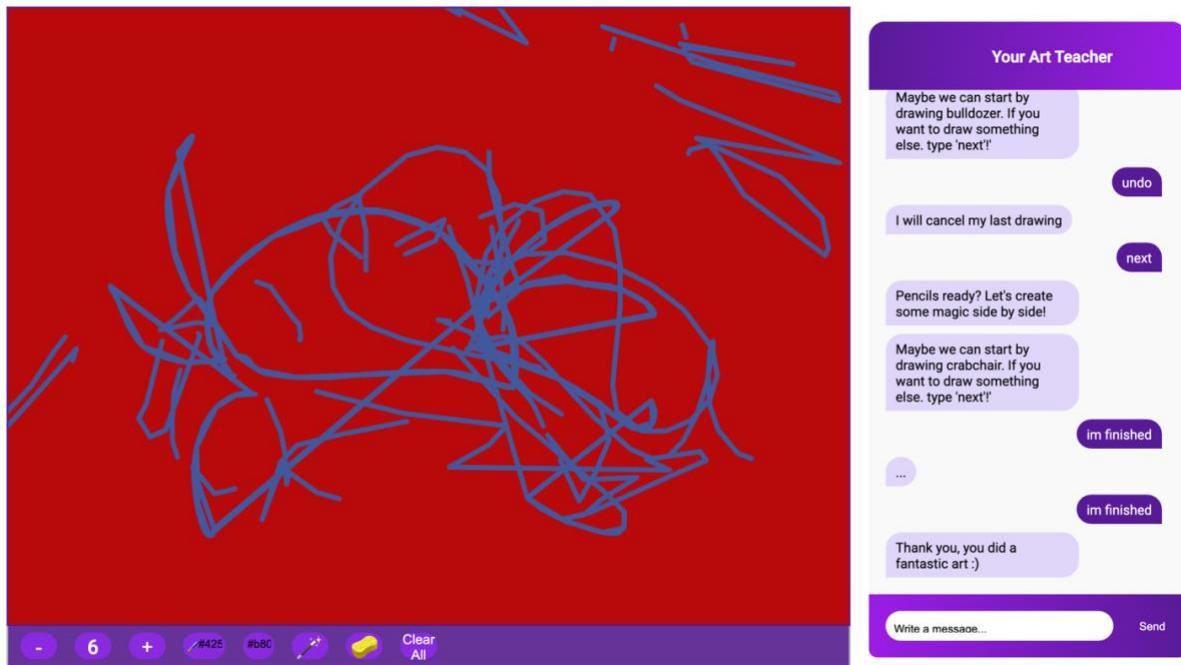


Image 7 - Example of Drawing Chatbot

The design of the ‘Drawing’ Chatbot feature was motivated by several key considerations:

1. Collaborative Creativity: research indicates that collaboration can enhance creativity by introducing new perspectives and ideas. The ‘Drawing’ Chatbot acts as a digital collaborator, adding elements to the drawing that the user might not have considered.
2. Direct Participation: By directly contributing to the drawing, the chatbot engages users in a shared creative process, making the activity more interactive and dynamic. This participation helps users feel supported and inspired throughout their creative journey.

By integrating these three distinct chatbots, each featuring specific AI capabilities, the application aimed to provide varied and enriching experiences, ultimately contributing to a deeper understanding of how AI-powered chatbots can influence and enhance creativity in digital drawing environments.

All programming codes have been uploaded to the GitHub¹⁹ repository for open access.

3.3 DATABASE INTEGRATION FOR COLLECTING DATA

In addition to user studies' questionnaires, it was decided to track and save user's activity during each application session, in case it correlates with the study results.

In all three versions of the application, I integrated an SQLite database. SQLite is a lightweight, serverless, and self-contained SQL database engine, widely used for its efficiency and ease of integration within applications. It requires minimal setup and provides solid performance, making it an ideal choice for tracking and storing user data in this study. The SQLite database tracked and saved various user interactions and activities, ensuring a comprehensive record of each session.

The specific data points collected included:

- User ID: A unique identifier for each user.
- Time Spent: The duration each user spent using the application, measured in seconds.
- Eraser Clicks: The number of times the user clicked the 'eraser' tool.
- Clear All Clicks: The number of times the user clicked the 'clear all' button.
- User Messages: The messages sent by the user during their session.
- Chatbot Messages: The messages sent by the chatbot in response to user interactions.
- Message Timestamps: The exact times when each message was sent, both by the user and the chatbot.
- Final Drawing Screenshot: A screenshot capturing the user's completed drawing at the end of their session.

All the data is automatically saved after the user sends a message, indicating that they are finished with their drawing. To manage and review the collected data, I implemented a Flask Admin panel. This panel displayed all the saved data in organised tables, allowing for easy access and oversight. The comprehensive dataset was further analysed to identify any correlations and insights relevant to the study's objectives. This detailed tracking and analysis were important for further understanding the impact of different chatbot features on user experience and creativity, ultimately contributing to the findings and conclusions of the study.

3.4 DESIGN OF THE USER STUDIES

To test the differences between the implemented chatbot versions, two user studies were conducted. The first study involved participants using each version of the app and filling out questionnaires afterward. The second study had participants ranking the creativity of the drawings produced in the first study.

3.4.1 FIRST USER STUDY

During the first user study, a consistent study environment was established for all participants. A dedicated room was booked, equipped with a laptop preloaded with the three app versions, and a Wacom Intuos Pro drawing tablet with a pen. The setup during the study is shown on the image 8:



Image 8 – The setup during the first user study

Participants were first required to sign a consent and data protection form. Following this, they completed a demographics questionnaire that included questions about their age, background, and experience with drawing and AI tools.

Participants were given a brief explanation of the study's goals and their tasks. They were allowed to practise drawing on the tablet for a few minutes to warm up. Once ready, they were asked to

draw anything they wanted and interact with the chatbot using each app version. The order of the app versions was varied to prevent order effects. After using each version for 10-15 minutes, participants completed two standardised questionnaires along with additional and qualitative questions. Qualitative questions included asking the specific features of the chatbot that either inspired or annoyed users and general opinion on the user experience of using the application. The quantitative questionnaire included two standardised assessments: the Creativity Support Index (CSI) and the Chatbot Usability Questionnaire. All answers were collected through using Google Forms.

3.4.1.1 Creativity Support Index (CSI)

The Creativity Support Index (CSI) was developed by Erin Cherry and Celine Latulipe to provide a standardised method for evaluating Creativity Support Tools (CSTs). The concept of creativity, while intuitively understood, is difficult to measure. Traditional methods for evaluating CSTs have relied heavily on qualitative approaches such as observations and interviews, which, while insightful, are often time-consuming, expensive, and challenging to compare across different studies [38].

To address these challenges, the CSI was introduced as a quantitative survey tool inspired by the NASA Task Load Index (TLX), which is widely used in the Human-Computer Interaction (HCI) community for assessing workload in productivity tasks. However, unlike the TLX, which measures factors such as time pressure and frustration, the CSI focuses on aspects more connected to creative work, such as exploration, expressiveness, and engagement. CSI consists of six key factors that contribute to creative work: Exploration, Collaboration, Engagement, Effort/Reward Trade-off, Tool Transparency, and Expressiveness [38]

The CSI has been employed in various contexts to evaluate digital tools designed to support creativity. For instance, Dewit et al. used the CSI to iteratively evaluate a Product-Service System (PSS) Design Toolkit in a graduate education setting. Over four years, the CSI provided insights that led to improvements in the toolkit, better supporting future designers in their creative challenges [39]

Participants rate their agreement with six statements, each representing one of the factors, on a Likert scale from 1 to 20. These statements are designed to capture different dimensions of the creative experience. Examples include:

- Exploration: "The tool encouraged me to explore many different ideas."
- Collaboration: "The tool supported collaboration with others."
- Engagement: "I felt deeply engaged while using the tool."
- Effort/Reward Trade-off: "The effort I put in was worth the results I achieved."
- Tool Transparency: "The tool was easy to use and did not get in the way of my creative process."
- Expressiveness: "The tool allowed me to express my ideas effectively."

In the second part, after using all three versions of the application, participants compare each of the six factors against each other to determine their relative importance. This part involves 15 paired comparisons (since each of the six factors is compared with every other factor), which helps identify which aspects of the tool were most importnat to the participant's creative process.

To calculate the overall CSI score, each factor's rating is multiplied by the count of times it was chosen as more important in the comparisons. These products are then summed and divided by 3, resulting in an index out of 100 [38].

3.4.1.2 Chatbot Usability Questionnaire (CUQ)

In order to evaluate how different features of chatbots affect the general user experience of communicating with the chatbot, The Chatbot Usability Questionnaire (CUQ) was used. CUQ was developed to provide a standardised method for assessing the usability of chatbots, recognizing that traditional usability scales, such as the System Usability Scale (SUS), are not fully suited to conversation-driven systems. The CUQ was created by an interdisciplinary team at Ulster University, aiming to measure various aspects of chatbot interactions, including chatbot personality, onboarding, user experience, and error handling [40]

The CUQ has been cited in multiple studies, demonstrating its effectiveness and adaptability in different domains. Its structured approach and validated metrics make it a valuable tool for both

academic research and practical evaluations of chatbot systems. The CUQ has been utilised in diverse research contexts to evaluate chatbot usability in fields such as healthcare, education, and customer service. For example, Holmes et al. (2019) applied the CUQ to assess the usability of a healthcare chatbot, finding that it provided reliable and valid results for evaluating conversational user interfaces [40] In educational settings, the CUQ helps researchers understand how well chatbots support learning activities and engagement. By using the CUQ, researchers and developers can identify strengths and areas for improvement in chatbot design, enhancing user experience and interaction quality across various applications.

The questionnaire consists of 16 both positive and negative statements, rated by participants on a Likert scale. Participants rate each statement on a scale from 1 to 5, where 1 indicates strong disagreement and 5 indicates strong agreement. The questions cover several dimensions of chatbot usability, such as:

- Chatbot Personality: Assessing the chatbot's ability to engage users in a human-like manner. Example: "The chatbot's responses were engaging and personable."
- Onboarding: Evaluating the ease with which users start interacting with the chatbot. Example: "It was easy to start a conversation with the chatbot."
- User Experience: Measuring overall satisfaction with the interaction. Example: "The chatbot provided helpful responses to my queries."
- Error Handling: Assessing the chatbot's ability to manage errors or misunderstandings. Example: "The chatbot handled misunderstandings effectively."

The scores from the CUQ are calculated similarly to the SUS. Each item's score contributes to an overall usability score out of 100. Positive responses (odd-numbered questions) are summed directly, while negative responses (even-numbered questions) are subtracted from the total to account for negative experiences. This method ensures a comprehensive assessment of both the strengths and weaknesses of the chatbot.

3.4.2 SECOND USER STUDY

The second user study was conducted to objectively assess the creativity of drawings produced by participants in the first user study. This additional study aimed to provide further validation or contradiction for the results obtained in the initial study and give more insights to them. A total of 17 individuals were recruited for this purpose. Participants did not know the details and objectives of this master's thesis to not make their rating biased. Participants were provided with a link to a Google Forms survey. The survey began with a demographics section, where participants answered questions about their age, educational background, and whether they worked or studied in an art-related field. This information was collected to understand the diversity of the participant pool and any potential biases related to their background.

Following the demographics section, participants were presented with sets of three drawings on each page. Each set included one drawing from each version of the chatbot created by a single user during the first study. Participants were asked to rate the creativity of each drawing on a scale from 1 to 7, where 1 indicated very uncreative and 7 indicated very creative. The example of such questions on creativity rating is show on image 9:

Image 9 - Example of CAT question

This comparative rating system is based on the Consensual Assessment Technique (CAT) [41], a standardised method widely used in creativity research to evaluate creative outputs.

3.4.2.1 Consensual Assessment Technique (CAT)

The Consensual Assessment Technique relies on the judgments of individuals who are familiar with the domain in question but do not need to be experts. The technique assumes that creativity is recognized through consensus among these informed judges. In this study, participants rated the creativity of each drawing relative to the other two in the set, ensuring that each drawing was evaluated in the context of its peers.

This rating method is based on the Consensual Assessment Technique (CAT), developed by Teresa Amabile. CAT is widely regarded as a reliable and valid method for evaluating creativity. It involves independent judges, who are familiar with the domain but not necessarily experts, rating the creativity of products relative to each other. The technique relies on the consensus among these judges to provide an objective measure of creativity. Over 40 years of research have validated CAT as a trustworthy method for creativity assessment across various domains, including art, writing, and problem-solving tasks [41].

CAT has been used extensively in creativity research. For example, Amabile's foundational work applied CAT to evaluate creative writing and artistic projects, demonstrating its effectiveness in capturing the nuanced judgments of creativity that standard tests often miss. It has also been employed in educational settings to assess the creativity of student projects and in organisational contexts to evaluate innovative products and processes [41].

By using CAT, this second study aimed to provide an objective measure of the creativity facilitated by each chatbot version. The creativity ratings obtained through this method offered additional validation for the subjective feedback from the first study. This combination of subjective and objective assessments provided a comprehensive understanding of the impact of different chatbot functionalities on user creativity.

3.5 STATISTICAL ANALYSIS METHODS

After collecting all data from the two user studies, the data was formatted, prepared, and cleaned in Excel tables for further analysis. The CSI and CUQ scores were calculated using their respective methods, with non-numerical answers (e.g., "agree", "disagree") converted to numerical values. Data from the database was also downloaded and organised into Excel tables. For the analysis, the JASP app¹⁹ was used. JASP (Jeffreys's Amazing Statistics Program) is an open-source statistical software that provides an easy-to-use interface for performing a variety of statistical tests.

The first step involved descriptive analysis where each questionnaire was described using mean values, standard deviation, and minimum and maximum values. This provided a basic understanding of the distribution and central tendencies of the data. Before starting comparative analysis, the normality of the distributions was tested using the Shapiro-Wilk test, along with visual inspections of histograms and QQ plots for CSI and CUQ results. These tests determined whether the data followed a normal distribution, which influenced the choice of subsequent statistical tests. Depending on the normality test results, either parametric or non-parametric tests were chosen. The decision also considered whether the independent variable was within-subject and if it was discrete.

Given the structure of the data and the study design, the Friedman test was selected as the most suitable statistical test. The Friedman test is a non-parametric test used to detect differences in treatments across multiple test attempts. It is particularly useful for comparing more than two related groups.

Effect size was calculated to determine the magnitude of the differences found. To identify specific differences between the chatbot versions, post hoc tests were applied. Post hoc tests are follow-up analyses used after finding significant results in the Friedman test, allowing for pairwise comparisons between groups.

19. JASP. Retrieved from <https://jasp-stats.org/>

Correlation tests were conducted to check if the results were related to any demographic factors or other variables. This analysis helped to identify potential influences on the creativity and usability scores.

Finally, the answers to the qualitative questions were thoroughly read and analysed for insights. This qualitative analysis provided context and depth to the quantitative findings, highlighting user experiences and perceptions that might not be captured through numerical data alone.

4. RESULTS

4.1 FIRST USER STUDY

4.1.1 DEMOGRAPHICS

We conducted the study with 12 respondents (6 males and 6 females) to collect information about their creativity practices, tool usage, and self-estimation on creativity. Their age was 23–31 years; the mean age was about 26 years, standard deviation (SD) = 1.9.

The study had several aspects of art creation and tool usage. Participants regularly engaged in drawing and painting, and the mean of the response equalled an "occasionally" rating ($M = 2.83$, $SD = 1.03$). High use of AI tools such as DALL-E, Midjourney, and ChatGPT was reported; the mean frequency was an "often" rating (mean = 4.00, $SD = 0.6$). However, when asked about the frequency of using AI tools for all forms of creative tasks, and specifically regarding drawing and sketching, the frequency became "occasionally" (mean = 1.67, $SD = 0.7$).

Digital drawing specifically was an unusual activity. The participants answered that they draw and sketch "occasionally" (mean = 1.75, $SD = 0.86$). The emotional and psychological aspects of art creation were also surveyed. Many people often experienced being stuck or uninspired while making art; the mean response was 3.80 and $SD = 1$, which is between "sometimes" and "often".

Self-assessment as to creativity indicated that participants felt creative in general; the average score was close to "often" (mean = 3.42, $SD = 0.5$). However, there is a sense of being uncreative at times; the average score in perceiving the lack of creativity in oneself was "sometimes" (mean = 3.00, $SD = 0.74$).

4.1.2 CSI RESULTS

In the first user study, participants rated three versions of an application, each with different functionalities of a chatbot—'Active', 'Recognizing', and 'Drawing'. Each participant used all three versions and completed the Creativity Support Index (CSI) questionnaire for each version to gauge

the support provided by the app for their creative processes. The results of CSI scores can be seen on the Figure 1:

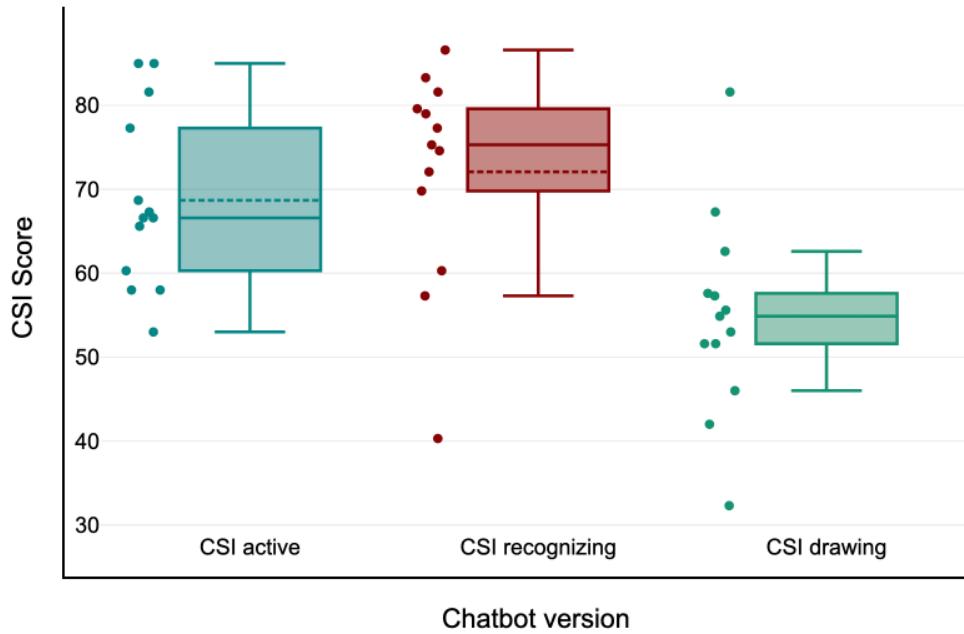


Figure1 - Results of CSI in a box plot

The mean CSI score for the 'Active' version of the app is 68.692, showing a moderate level of perceived support, where the lowest score was recorded as 53.00 and the highest at 85.00. The standard deviation of 11.03 suggests the group is made up of a diverse set of responses among the participants.

The 'recognizing' version scored a bit higher on the CSI, with a mean of 72.083. This app version seemed most effective, with the highest median and maximum score at 86.600 but with the biggest spread of scores, as indicated by the standard deviation of 13.308. This shows diverse levels of perceived effectiveness across the participants.

On the other hand, the 'drawing' version received the lowest mean CSI score = 54.875. It showed scores ranging from 32.300 to 81.600 and a standard deviation of 12.509, showing significant variability in the support of creativity by that version.

4.1.3 CUQ RESULTS

The same group of 12 participants rated each version of the app—'Active', 'Recognizing', and 'Drawing'—on the Chatbot Usability Questionnaire (CUQ). This rating provided an assessment of the various usability aspects for each chatbot functionality among the same group of 12 participants. The results of CUQ scores can be seen on the Figure 2:

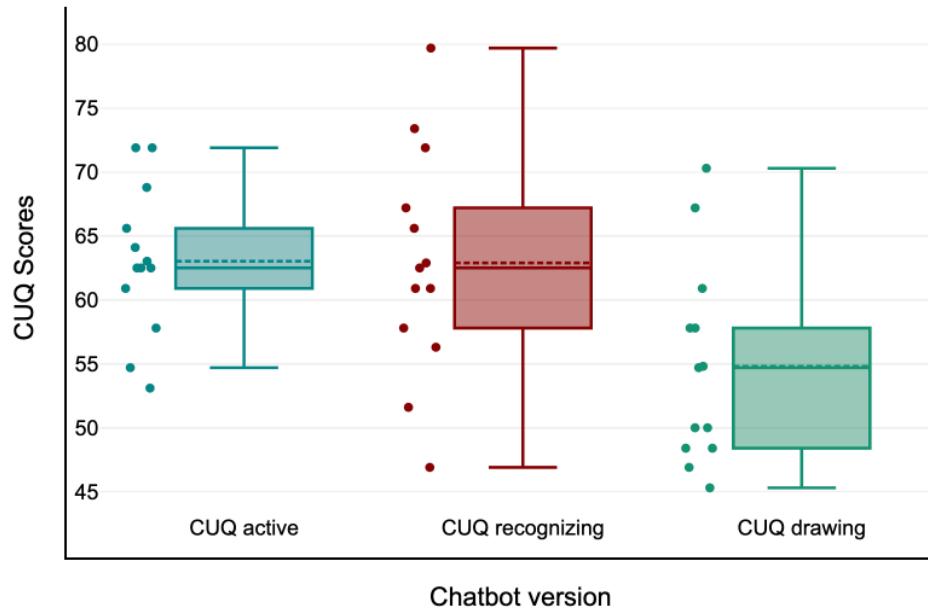


Figure 2 - Results of CUQ in a box plot

The 'Active' version of the chatbot showed a mean usability score of 63.025. It was the most consistent version to receive usability evaluations in that it had the least variability in responses, indicated by a standard deviation of 6.024. The 'Active' version's scores ranged from a minimum of 53.100 to a maximum of 71.900—generally showing highly favourable usability perceptions.

The 'recognizing' version of the chatbot received a slightly lower mean score of 62.892. Although it had a maximum score of 79.700—indicating some users found it very usable—the wide dispersion of scores in the standard deviation of 9.345 implies relatively wider experience among participants, with the lowest score reaching 46.900.

The 'drawing' version of the chatbot has the lowest usability scores. This version's mean score was 54.808, indicating less favourable usability perceptions. Its standard deviation of 8.146 and scores

ranging from 45.300 to 70.300 again point to great variability in how participants perceived the usability of this version.

4.1.4 USER INTERACTION

The study also examined the duration of use and frequency of specific actions used by users while interacting with each different version of the app such as using the eraser, deleting all content, and sending messages. This data provides insight into how participants interacted differently with each version of the application

The time that participants spent on the 'Active' and 'Recognizing' versions of the app was almost identical with mean times of approximately 5.808 minutes ($SD = 2.4$) and 5.792 minutes ($SD = 2.5$), respectively. The 'Drawing' version was used for a slightly shorter mean time of 5.65 minutes ($SD = 2.8$), but still almost for the same time. The standard deviations suggest moderate spread across sessions, especially in the 'Drawing' version where the spread between the minimum (3.1 minutes) and maximum (13.3 minutes) usage time was the widest.

The data save in database also provides the following regarding 'eraser' button usage:

- 'Active' version: Used with low frequency on the mean (2.0 times, $SD = 3.7$), ranging between 0 to 11 times.
- 'Recognizing' version: More frequently used on the mean (3.0 times, $SD = 4.1$), also showing a wide range between 0 to 12 times.
- 'Drawing' version: Least frequently used on the mean (0.5 times, $SD = 0.8$), with the minimum variance.

'Clear All' button usage:

- 'Active' version: Rarely used on the mean (1.1 times, $SD = 2.6$), meaning few sessions needed a total restart.
- 'Recognizing' version: Similarly low use on the mean (2.4 times, $SD = 3.1$).
- 'Drawing' version: Used more often on the mean (5.1 times, $SD = 8.9$), implying more trial and error or dissatisfaction with results.

Number of messages sent to Chatbot:

- 'Active' version: Fewer messages sent on the mean (4.58, SD = 2.3), possibly suggesting a more straightforward or less interactive feature.
- 'Recognizing' version: A moderate number of messages sent on the mean (5.75, SD = 4.6), indicating a higher level of interaction.
- 'Drawing' version: Most messages were sent on the (mean = 8.75, SD = 4), indicating the highest level of interaction or need for more guidance and feedback from the chatbot.

4.1.5 COMPARATIVE ANALYSIS

A statistical analysis of the user experience of the three versions of the application—‘Active’, ‘Recognizing’, and ‘Drawing’—was analysed using the Creativity Support Index and the Chatbot Usability Questionnaire. With the initial assumption, the results of the Shapiro-Wilk test indicated that the data were normally distributed. However, after visualising histograms and QQ plots, the distribution appeared to be non-normal. Therefore, nonparametric tests were performed for further analysis.

4.1.5.1 CSI

The Friedman test, a nonparametric replacement for repeated measures ANOVA, was conducted to compare CSI scores across the three versions of the application. The test statistic was significant ($\chi^2 = 13.1667$) with a p value of 0.001383. The effect size was also calculated: $W = 0.55$, which means a large difference between the ranks of versions, which confirms the practical significance of the results obtained. The results of Friedman test are shown on the Figure 3:

Factor	Chi-Squared	df	p	Kendall's W
Prototype	13.167	2	0.001	0.549

Figure 3 - Friedman test for CSI results

Post hoc differences were also further tested, and they are shown on the Figure 4:

		Mean Difference	SE	t	p _{holm}
Active	Recognizing	-3.392	3.367	-1.007	0.325
	Drawing	13.817	3.367	4.104	< .001
Recognizing	Drawing	17.208	3.367	5.111	< .001

Figure 4 - Post Hoc tests for CSI results

The results are:

- Active vs. Recognizing: A p-value of 0.325 states that there is no significant difference between those two versions. The mean difference and Cohen's d value also support the conclusion of minimal practical difference between the two versions.
- Active vs. Drawing: The mean difference of -13.817 with a significant p-value (<0.001) and Cohen's d of -1.122 suggests that the 'Active' version is significantly preferred or performs better than the 'Drawing' version.
- Recognizing vs. Drawing: Similarly, the mean difference of -17.208, with a p-value <0.001 and Cohen's d of -1.397, indicates a significant preference for the 'Recognizing' version over the 'Drawing' version.

4.1.5.2 CUQ

The Friedman test shows a very large value of chi-squared of 4.638 in 2 degrees of freedom with a corresponding p-value of 0.098. The p-value received is not statistically significant enough to confirm a statistically meaningful difference in the CUQ scores of the three versions of the application and the conventional alpha level of 0.05. The effect size is presented as Kendall's W at 0.193, which is a moderate level of agreement among the rankings of the three versions. The results of Friedman test are shown on Figure 5:

Factor	Chi-Squared	df	p	Kendall's W
Versions	4.638	2	0.098	0.193

Figure 5 - Friedman test for CUQ results

A more detailed comparison among the pairs of app versions is presented by Post Hoc tests on Figure 6:

		Mean Difference	SE	t	Cohen's d	p_{holm}
Active	Recognizing	0.133	3.326	0.040	0.017	0.968
	Drawing	8.217	3.326	2.470	1.033	0.065
Recognizing	Drawing	8.083	3.326	2.430	1.016	0.065

Figure 6 - Post Hoc tests for CUQ results

The description of results:

- Active vs. Recognizing: Mean Difference = 0.133, SE = 3.326, t = 0.040, d = 0.017, $P_{\text{holm}} = 0.968$. As can be seen, there is no significant difference between these two versions with respect to their usability as rated by CUQ.
- Active vs. Drawing: Mean Difference = 8.217, SE = 3.326, t = 2.470, d = 1.033, $P_{\text{holm}} = 0.065$. The result points in the direction of significance, which, however, has not yet crossed the conventional threshold of 0.05. The comparison somehow hints at a meaningful difference in favour of the 'Active' version, but it is not significant.
- Recognizing vs. Drawing: Mean Difference = 8.083, with SE of 3.326 and a similar t-value of 2.430. The effect size (d) here is substantial at 1.016, and the adjusted p-value is again 0.065. As in the previous comparison, this again suggests a statistically meaningful but not definitive preference for the 'Recognizing' version compared to the 'Drawing' version.

4.1.6 CORRELATIONS

In an effort to uncover what might be influencing the results of the study, a detailed correlation analysis using Spearman's Correlation test was made. This examination looked at the interplay between various demographic factors—like age, gender, participants' responses to certain questions and their actions—and how these might impact the effectiveness of the app, as measured by the Creativity Support Index (CSI) and the Chatbot Usability Questionnaire (CUQ).

However, the correlation results revealed that most of these demographic and interaction factors didn't really affect the CSI and CUQ outcomes. This also held true for the analytics pulled from

the app database: how long participants used the app, how often they used the eraser or delete functions, or how many messages they sent. None of these seemed to correlate with the users' perception of the app's support for creativity or its usability. However, a couple of correlations stood out:

- Users who often create digital art felt more supported by the 'Active' version of the app: Specifically, those who regularly engage in digital art creation tended to rate the 'Active' version higher on the CSI. The strength of this relationship was strong, with a correlation coefficient of .664 and a p-value of .019, underscoring a meaningful connection between frequent digital art activity and a positive experience with the 'Active' version.
- Frequent digital art creators also preferred the usability of the 'Drawing' version: Similarly, those who often create digital art found the 'Drawing' version of the app to be more user-friendly, reflected by a correlation coefficient of .681 and a p-value of .015.

Additionally, the correlations between versions ordering with CSI and CUQ results were made, but they didn't reveal any significant results.

4.1.7 QUALITATIVE ANALYSIS

Additionally to the CSI questionnaire and CUQ, several qualitative questions were asked after each version of the application, as well as a couple questions at the end of the study. After completing using each version of the app, users were asked to answer which features of the app made them feel inspired and creative, which features made them feel annoyed and distracted and whether they felt stuck/uninspired during the process of drawing. Based on the answers to those questions, the following patterns and distinctions were found:

Inspirational and Creative Features:

- Active Version: Users appreciated the need for structured guidance, relevance, and autonomy when implementing the suggestions of the chatbot. This version seemed to set up a supportive space by providing concrete steps and advice. Some examples of responses that users provided were:
 - “I like how it gave me steps of how to draw”

- “Yes, adding shadows was a good idea”
 - “I gave it a prompt for a funny picture idea, and it did its job well”
- Recognizing Version: The guessing feature and friendly interaction were appreciated by users for inspiring creativity. Some example of responses that users provided were:
 - “I like how the behaviour was friendly with emojis, and was interesting to see what the chatbot guesses what I am drawing.”
 - “Even though it recognized my ideas in different way, but giving me some interesting points which is great to make things creative”
- Drawing Version: This version was appreciated for being interactive and somewhat unconventional in the suggestions made by the chatbot, such as adding random elements that evoke abstract thinking. Users enjoyed seeing how the chatbot could interact with their drawings. Some example of responses that users provided were:
 - “It was interactive.”
 - “It added stuff which seemed pretty random but which inspired me to include other things in my drawing.”

Annoying or Distracting Features:

- Active Version: Main issues include overactivity - the frequency and length of messages sometimes distracted users from the creative process. Some example of responses that users provided were:
 - “I found the messages useful but they came up a little bit too often”
 - “The paragraphs were too long at times.”
- Recognizing Version: Similar to the Active version, too frequent guessing and sometimes irrelevant messages were distractions for some users. Some example of responses that users provided were:
 - “it answered way too much. the responses were somehow irrelevant or wrong.”
 - “While it recognized correctly (which I told them) and I was just adding colour, it continued guessing which I found confusing.”
- Drawing Version: The major point of contention was that the chatbot would draw over users' work without apparent intention, which frustrated some users. The randomness of

these additions was usually seen as a distraction rather than an inspiration. Some examples of responses that users provided were:

- “Yes. It usually added random paint and lines onto my drawing which made no sense. It felt like random additions.”.
- “The drawings of the chatbot made the drawings pointless because it just drew weird lines over the drawing.”

Moments of Feeling Stuck or Uninspired:

- Active Version: Several users felt stuck, some due to not receiving specific tips from the chatbot. But the majority didn't have any issues with feeling uninspired and stuck without any further explanation.
- Recognizing Version: Some users said the chatbot provided good starting points to decrease the feeling of being stuck, while only a couple felt stuck or uninspired.
- Drawing Version: Some users noted that they were forced to restart their work due to added random lines, but in general they found these features exciting without feeling stuck and uninspired.

Perception of chatbot interface in the digital drawing environment:

- Most of users found it appropriate by sending messages such as: “I think it was appropriate”, “I believe it is a fun way to draw digitally especially for beginners so that users can have some other interaction and have more ideas what they can do. “, “I really like the idea, it has a lot of potential in the field of educational AI apps. Considering the amount of time children spend on tablets every day for entertainment, it can be a useful tool for a creative and educational development of kids.”

4.1.8 ADDITIONAL ANALYSIS

Also, to support standardised CSI and CUQ questionnaires, three additional questions were asked after each version to find out whether the chatbot of a specific version of the application inspired new ideas, engaged in the drawing process or distracted from the drawing process. Answers were

collected and further analysed. Users were asked to answer to the following question by 5-point Likert scale:

Q1: Chatbot inspired me to new ideas for drawings

- Active Version: Shows positive feedback with a mean of 3.833 (SD = 0.6), suggesting that most users agreed that the chatbot inspired new drawing ideas.
- Recognizing Version: Slightly lower inspiration with a mean of 3.583 (SD = 0.9), also indicating general agreement.
- Drawing Version: Similarly positive with a mean of 3.667 (SD = 1.1), indicating that it was also effective at inspiring users, almost as the Active version.

Q2: Chatbot messages kept me engaged in the drawing process

- Active Version: Moderate engagement with a mean of 3.583 (SD = 0.79). This suggests that users generally agreed that the chatbot kept them engaged.
- Recognizing Version: Slightly lower engagement with a mean of 3.417 (SD = 1.0), indicating agreement but less compared to the Active version.
- Drawing Version: The least engaging with a mean of 3.083 (SD = 1.0), which leans towards neutral, suggesting that some users felt the messages did not maintain their engagement effectively.

Q3: Chatbot distracted me from the drawing process

- Active Version: Indicated minimal distraction with a mean of 2.417 (SD = 0.79), mostly disagreeing with the statement that it was distracting.
- Recognizing Version: Also showed low distraction levels with a mean of 2.417 (SD = 0.79), same as the Active version.
- Drawing Version: Higher distraction level with a mean of 3.083 (SD = 1.0), leaning closer to neutral but tending towards agreement, indicating this version caused more interruptions.

Due to the data being normally distributed, the comparative analysis of the collected data was made by using Repeated Measures ANOVA test, which revealed no significant results.. Moreover, at the end of the user study participants were also asked to choose which version of the application they like the most and the least. According to their answers, the majority chose the ‘Recognizing’

version as the favourable one, while the 'Drawing' one was the least liked one, which corresponds to the results of CSI scores.

4.2 SECOND USER STUDY

4.2.1 DEMOGRAPHICS

In a second user study involving the Consensual Assessment Technique, 17 participants (3 males, 13 females, and 1 non-binary) evaluated the creativity of sketches created by users from an initial study. The participants varied widely in their educational backgrounds, holding qualifications ranging from high school diplomas to Master's degrees in diverse fields including computer science, international relations, and chemistry.

The demographic profile of the participants was as follows:

- Age: The ages of the participants ranged from 17 to 30 years, with an average age of 25.41 years and standard deviation = 3.08.
- Experience in an Art-Related Field: About 47.1% of participants reported having work experience in an art-related field.
- Education in an Art-Related Field: Less than a third of the participants (29.4%) had studied in an art-related field, suggesting a minority had formal education in arts.
- Frequency of Creating Art: On average, participants engaged in art creation like drawing or painting with a mean score of 2.529 ($SD = 1.13$), which corresponds to "occasionally." The range of responses varied from "never" (1) to "always" (5).

4.2.2 CAT RESULTS

All 17 participants rated the creativity of sketches created using three different versions of a chatbot. Each participant rated sketches on a scale from 1 (very uncreative) to 7 (very creative), comparing sketches produced by the same user to ensure a consistent basis for comparison. The mean values of each participant's ratings on each version's drawings were calculated and used for the further analysis. The results can be seen on Figure 7:

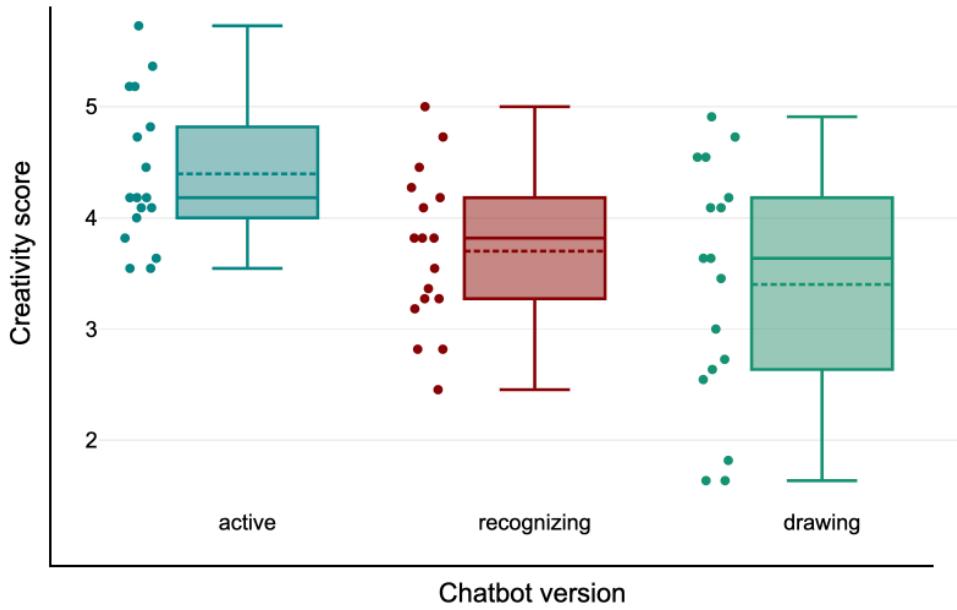


Figure 7 – Mean creativity scores for 3 versions of the app by each participant

The following information was analysed:

- The ratings for the Active version varied from a low of 3.545 to a high of 5.727, with a standard deviation of 0.665 with the mean value of 4.396, indicating relatively consistent ratings around the central tendency.
- Scores for the Recognizing version were spread from 2.455 to 5.000, with a standard deviation of 0.705 and the mean value of 3.701. This spread shows a bit more variability in how participants perceived the creativity of sketches compared to the Active version.
- The ratings for the Drawing version showed the widest range, from 1.636 to 4.909, and had the highest standard deviation of 1.090. The mean value is 3.401. This suggests a greater divergence in participant opinions regarding the creativity level in the sketches from the Drawing version.

4.2.3 COMPARATIVE ANALYSIS

The result from the second user study was statistically analysed by also using the Friedman test for estimating the difference in creativity ratings between three versions of the app. This was followed by Post Hoc testing in order to establish differences between each version.

The Friedman test showed that there was a statistically significant difference in the creativity ratings between the three versions of the app, with a p-value of 0.004. This means that at least one of the versions of the app was rated differently in terms of perceived creativity compared to the others. The Kendall's W of 0.327 hints at a moderate effect size, which means that the difference in ratings has practical significance. The results of Friedman test are shown on Figure 8:

Factor	Chi-Squared	df	p	Kendall's W
Prototype	11.121	2	0.004	0.327

Figure 8 - Friedman test for Consensual Assessment Technique results

To find out the difference between versions Post Hoc testing was made shown on Figure 8:

		Mean Difference	SE	t	p _{holm}
Active	Recognize	0.695	0.267	2.607	0.028*
	Drawing	0.995	0.267	3.730	0.002**
Recognize	Drawing	0.299	0.267	1.123	0.270

Figure 8 - Post Hoc Tests for creativity ratings in the second user study

The results of Post Hoc test are:

- Active vs. Recognizing: The Post Hoc test reveals that there is a statistically significant difference between 'Active' and 'Recognizing' versions. The adjusted p-value of 0.028 confirms that this result is reliable at a 5% significance level.
- Active vs. Drawing: There is also a statistically significant difference between the 'Active' and 'Drawing' versions. The adjusted p-value of 0.002 indicates a highly significant difference. This suggests a strong preference or perception of greater creativity in the 'Active' version over the 'Drawing' version.

- Recognizing vs. Drawing: In contrast, there is no statistically significant difference in the creativity ratings between the 'Recognizing' and 'Drawing' versions (p-value is 0.270)

Worth noting, the results of the Pearson correlation indicated that there is a non-significant medium positive relationship between participants' experience in art creation and their creativity ratings on "Drawing" version, ($r(15) = .317$, $p = .216$). The correlations analysis between the other demographics factors and creativity scores did not reveal any significantly important results.

5. DISCUSSION

This study aimed to evaluate the impact of different AI-enhanced chatbot features within a digital drawing application, focusing on two primary research objectives: evaluating the influence of chatbot features on supporting users' creativity and the overall user experience. The data from the demographic's questionnaire indicates that while the use of AI tools is frequent for art creation, it is less common for more specific tasks like digital drawing and sketching. Additionally, there is a notable frequency of feeling uninspired during the creative process, which could influence how users perceive their own creativity. Therefore, the findings from both user studies provided insights into the study objectives, which can be valuable for future implementations.

5.1 ASSESSING THE INFLUENCE OF CHATBOT FEATURES ON SUPPORTING USERS' CREATIVITY

The first research objective was to determine how various chatbot functionalities support users' creativity while using a digital drawing application. The results showed significant differences in the effectiveness of the three chatbot versions in supporting creativity. The comparative analysis, supported by statistical tests of significance, indicated that in terms of CSI there was no significant difference between the "Active" and "Recognizing" versions, but both versions were significantly more preferred than the "Drawing" version. Therefore, it is important to discuss features of each chatbot version to understand their influence on the user's creativity.

The 'Recognizing' version received the highest mean CSI score and showed significant preference in results, indicating that users felt most creatively supported and engaged by the chatbot's ability to guess their drawings. According to the data from the database, users sent a moderate number of messages to the chatbot and a lot of them were messages such as 'what am I drawing now?', 'what did I draw?', which indicated users' interest in the chatbot's recognizing feature. Based on the guesses from a chatbot, users were either inspired to change their drawings (by sending messages like 'should I add something?') or start over and draw a new subject or redraw their initial sketch more accurately. Therefore, it can be assumed that recognizing feature of the chatbot likely made users feel acknowledged and encouraged them to experiment, validating their creativity and exploring new ideas. The proactive interaction of the 'Recognizing' chatbot seemed to strike a

balance between offering support and allowing users the freedom to explore their creativity without feeling overwhelmed.

The 'Active' version had a slightly lower mean CSI, but also showed significant preference in results. Users appreciated the continuous support from initiation and the engaging messages from the chatbot. Although in this version users had the least mean number of messages sent, messages were the most profound - users asked for specific advice (example: "can you give me some advice when it comes to urban drawings", "give me idea for a funny sketch") and reacted to the general advices from the chatbot (example: "How did you notice I didn't have my back straight"). Users also had the least number of 'Clear All' button clicks, which implies that they could be more focused on details of producing one drawing. However, some participants found the frequency and verbosity of the messages distracting, which may have detracted from their creative process. Despite this, the 'Active' version was effective in maintaining user engagement and providing a supportive environment for creativity.

The 'Drawing' version, on the other hand, received the lowest CSI score and was significantly less preferred. This version's feature of directly participating in the drawing process by adding lines to the artwork was perceived as more intrusive than supportive. Although this version had the highest mean number of the messages sent by users, most of them were just instructing chatbot to either undo its additions or draw the next object. Users also used the 'Clear All' button quite frequently, which implies that they couldn't focus on one drawing, as well as that they were trying out different objects to draw. While some users enjoyed the collaborative aspect and found it novel and inspiring, others felt frustrated by the direct additions of lines, which often disrupted their creative flow. This indicates that while collaborative creativity can be beneficial, it must be carefully balanced to avoid interfering with the user's autonomy and creative control.

The second user study, using the Consensual Assessment Technique (CAT), provided an objective measure of the creativity in the drawings produced using each version. From the comparative analysis results, the 'Active' version of the app is seen as significantly more creative compared to both the 'Recognizing' and 'Drawing' versions. There is, however, no significant difference in creativity perception between the 'Recognizing' and 'Drawing' versions. The results reinforce the

findings from the first study, with the 'Active' version's drawings rated as the most creative, and the 'Drawing' version's drawings as the least. However, the results do not align for the 'Recognizing' chatbot, which implies that although participants of the first study felt more creatively supported by that version, the drawings they produced were not objectively perceived as more creative. This indicates the need for further research in how the recognizing features of chatbots objectively affect users' creativity.

5.2 EVALUATING USER EXPERIENCE

The second research objective focused on the overall user experience with different chatbot features. The Chatbot Usability Questionnaire (CUQ) provided insights into the usability aspects of each version. The comparative analysis did not indicate overall substantial differences across the three versions, it only pointed on some tendency for preferring the 'Active' and 'Recognizing' versions over the 'Drawing' version, although it did not reach the conventional threshold of statistical significance. The 'Active' and 'Recognizing' versions received similar CUQ scores, indicating that both were relatively well-received in terms of usability. Users found these versions engaging and appreciated the relevant and timely feedback provided by the chatbots. However, the 'Drawing' version lagged behind, with lower CUQ scores reflecting users' dissatisfaction with its usability. The primary issue was the perceived intrusiveness of the chatbot's contributions to the drawing, which many users found distracting rather than helpful. This highlights the importance of ensuring that AI-enhanced features support rather than dominate the creative process.

Interesting to note, although the correlation analysis didn't reveal that much information, it still pointed out that users who frequently engaged in digital art creation felt more supported by the 'Active' version and found the 'Drawing' version more user-friendly. This suggests that experienced digital artists may have specific expectations and preferences that influence their perception of chatbot functionalities. It can also be assumed that they were more interested in collaboration and intrusiveness of the 'Drawing' chatbot and in initiative behaviour of 'Active' chatbot due to its novelty compared to their usual experience with digital drawing. The assumptions regarding 'Drawing' version can also be supported by a medium positive correlation between the participants' experience in art creation and their creativity ratings in CAT, although it is not

statistically significant. Therefore, further research is needed to approve or disapprove such insights.

5.3 LIMITATIONS OF THE STUDY

One crucial limitation was the restricted availability of drawing APIs. The development of the chatbot features relied on a limited selection of APIs, which constrained the performance of the application. The APIs used were not only limited in number but also showed imperfections that occasionally affected the user experience and the accuracy of the chatbot's responses. For instance, the QuickDraw API sometimes made errors in recognizing user input, which could lead to inaccurate feedback or suggestions from the chatbot. These errors likely diminished the overall effectiveness of the feature and contributed to user frustration. Similarly, the 'Drawing' version's reliance on SketchRNN API to add random elements to user drawings also faced challenges. The randomness and occasional lack of relevance of these additions were partly due to the limitations of the APIs in generating contextually appropriate suggestions.

Despite the insightful findings, user studies also had several limitations that need to be addressed:

- Firstly, the sample size was relatively small, with only 12 participants in the first user study and 17 in the second. While the results provide initial insights, a larger sample size would be necessary to generalise the findings more broadly.
- Secondly, the demographic diversity of the participants was limited. Most participants were young adults with varying degrees of familiarity with digital drawing and AI tools. Future studies should aim to include a more diverse age group and individuals with different levels of expertise in digital art to better understand the impact of chatbot features across a broader population.
- Thirdly, the study was conducted in a controlled environment with a specific setup (laptop and drawing tablet). While this ensured consistency, it may not fully capture the varied contexts in which users engage with digital drawing applications. Exploring the impact of chatbot features in more naturalistic settings, such as on different devices and in different environments, would provide a more comprehensive understanding of their effectiveness.

- Finally, the study focused on the immediate impact of chatbot features on creativity and user experience. Long-term studies are needed to understand how these features influence creativity over extended periods and how users' perceptions and interactions with the chatbots evolve over time.

While this study provides valuable insights into the role of AI-enhanced chatbots in supporting creativity in digital drawing applications, addressing these limitations in future research will be crucial for developing more effective and user-friendly creativity support tools.

6. CONCLUSION

The exploration into the integration of AI-enhanced chatbots within digital drawing applications highlights important insights into how different chatbot features can influence creativity and user experience. The study was motivated by the increasing use of AI tools in creative fields and aimed to fill the gap in understanding how specific chatbot functionalities affect users' creative processes and overall satisfaction.

The research involved developing a web-based drawing application incorporating three distinct versions of AI-enhanced chatbots, each designed with unique features: 'active,' 'recognizing,' and 'drawing.' Through two comprehensive user studies, the impact of these chatbots on creativity and user experience was meticulously evaluated.

6.1 KEY FINDINGS

The main findings regarding chatbot version are the following:

1. 'Active' Chatbot:

- The 'active' chatbot provided continuous support and engagement by offering drawing tips and initiating frequent interactions.
- Users appreciated the structured guidance, which helped maintain their focus and prevent creative blocks.
- However, the frequency and verbosity of messages were sometimes perceived as distracting, indicating the need for a balance between proactive engagement and user autonomy.

2. 'Recognizing' Chatbot:

- The 'recognizing' chatbot, which identified users' drawings and provided feedback, received the highest Creativity Support Index (CSI) score.
- This feature made users feel acknowledged and encouraged experimentation, validating their creative efforts and inspiring new ideas.
- The dynamic interaction and balance between support and freedom contributed to its effectiveness.

3. 'Drawing' Chatbot:

- The 'drawing' chatbot, which actively participated in the drawing process by adding lines, was the least preferred in terms of CSI and user experience.
- While some users enjoyed the collaborative aspect, many found the random additions intrusive and disruptive to their creative flow.
- This highlights the importance of careful implementation of collaborative features to avoid interfering with users' creative control.

Objective Creativity Assessment:

- The second user study, using the Consensual Assessment Technique (CAT), confirmed that drawings produced with the 'active' chatbot were rated as the most creative.
- There was no significant difference in perceived creativity between the 'recognizing' and 'drawing' versions, despite the 'recognizing' chatbot being preferred in terms of CSI.

User Experience:

- The Chatbot Usability Questionnaire (CUQ) results indicated that while there were tendencies favouring the 'active' and 'recognizing' versions, the differences were not statistically significant.
- Users' qualitative feedback revealed that inspirational features, dynamic engagement, and the balance of interaction played crucial roles in their overall experience.

As a result, we can state that the first hypothesis (H1) of the study objectives, which posited that different chatbot features would significantly support users' creativity to varying degrees was supported by the results and the null hypothesis is strongly rejected. The second hypothesis (H2) suggested that chatbots with different features would result in significantly different user experience levels. This hypothesis was partially supported, but the null hypothesis can not be strongly rejected; although the Friedman test for Chatbot Usability Questionnaire (CUQ) scores did not show a statistically significant difference at the conventional alpha level, qualitative feedback and post hoc tests indicated a trend towards a preference for the 'Active' and 'Recognizing' versions over the 'Drawing' version. Thus, while the user experience differences

were not as pronounced as the creativity support differences, there was still a discernible pattern favouring certain chatbot features over others.

The study underscores the potential of AI-enhanced chatbots to notably support and enhance creativity in digital drawing applications. It demonstrates that while proactive and engaging chatbots can foster creativity, the implementation must be balanced to avoid overwhelming users. The findings provide valuable insights for designing more effective AI-enhanced creativity support systems, emphasising the need for features that inspire without intruding, engage without distracting, and support without taking control. Future research should continue exploring these dynamics to optimise the integration of AI in creative domains, ensuring that these tools augment rather than hinder the creative process.

6.2 FUTURE DESIGN OF AI-ENHANCED CHATBOTS IN DIGITAL DRAWING

The findings from this study have substantial implications for the design of AI-enhanced chatbots in digital drawing applications. Users generally find the chatbot interface in digital drawing applications appropriate and enjoyable, especially for beginners and educational purposes. Several respondents highlighted its potential in educational AI apps for children. The main outcome of the study's findings which could be beneficial for future design of chatbot in a digital drawing environment:

- To effectively support creativity, chatbots should provide relevant and timely feedback without overwhelming the user, but still be initiative and active.
- Features that recognize and validate user input, as seen in the 'Recognizing' version, can enhance the creative process by making users feel acknowledged and inspired. However, developers should consider incorporating more advanced image recognition for better results.
- Collaborative features, such as those in the 'Drawing' version, need to be carefully designed to ensure they complement rather than disrupt the user's creative flow, possibly through real-time adaptation and user control mechanisms.

Moreover, qualitative feedback indicated that some features of the chatbots were both appreciated and criticised by different users. For instance, while the 'Recognizing' chatbot guessing feature was engaging for most users, it was seen as distracting by some others. This suggests that user preferences for certain features can be highly individual. Future research should consider more personalised approaches to chatbot design, allowing users to customise the level and type of interaction they receive based on their preferences and needs. Also based on the results, it might be beneficial to adapt different chatbot features in application for variously skilled users, since some features such as collaborative drawing seems to be more appreciated by users with some experience in digital drawing.

In conclusion, the study shows that AI-enhanced chatbots have great potential to support and enhance creative skills in digital drawing applications. Thus, the research findings provide essential insights for designing more efficient AI-enhanced creativity support systems with chatbots, highlighting the importance of features that inspire without intruding, engage without distracting, and support without taking control. Future studies should further explore these dynamics to optimise the integration of AI chatbots in creative domains, ensuring that these tools enhance rather than hinder the creative process. This means that it can also be used as a reference for other AI-enhanced chatbots in other creative domains, in order to support and enhance creativity in different applications.

7. REFERENCES

- [1]. Laavanya Ramaul, Paavo Ritala, and Mika Ruokonen. 2024. Creational and conversational AI affordances: How the new breed of chatbots are revolutionizing the knowledge industries. *Business Horizons*. <https://doi.org/10.1016/j.bushor.2024.05.006>
- [2]. M. A. Runco and G. J. Jaeger. 2012. The standard definition of creativity. *Creativity Research Journal*, 24(1), 92-96. <https://doi.org/10.1080/10400419.2012.650092>
- [3]. C. Heck. 2017. The Rise of Digital Art. The Broadcaster. <https://hhsbroadcaster.com/4399/ane/the-rise-of-digital-art>
- [4]. R. K. Sawyer. 2012. Explaining creativity: The science of human innovation. Oxford University Press. <https://doi.org/10.1093/oso/9780195161649.001.0001>
- [5]. Benedek, M., Jauk, E., Fink, A., Koschutnig, K., Reishofer, G., Ebner, F., & Neubauer, A. C. (2014). To create or to recall? Neural mechanisms underlying the generation of creative new ideas. *NeuroImage*, 88, 125-133. <https://doi.org/10.1016/j.neuroimage.2013.11.021>
- [6]. Drake, J. E., & Winner, E. (2018). Expression Versus Distraction: An Investigation of Contrasting Emotion Regulation Strategies When Drawing, Writing, Talking, and Thinking. *Empirical Studies of the Arts*, 36(2), 162-179. <https://doi.org/10.1177/0276237417718>
- [7]. C. Paul. 2023. Digital Art (World of Art). Thames & Hudson.
- [8]. M. Zagrobelna. 2017. How Is Digital Drawing Different from Traditional Art? Medium. <https://medium.com/sketchbook-in-perspective/how-is-digital-drawing-different-from-traditional-art-8204e5954d6b>
- [9]. C. H. Faber. 2009. Digital Drawing Tablet to Traditional Drawing on Paper. Iowa State University, College of Design.
- [10]. Shneiderman, B. (2007). Creativity support tools: Accelerating discovery and innovation. *Communications of the ACM*, 50(12), 20-32. <https://doi.org/10.1145/1323688.1323689>
- [11]. J. McCormack, T. Gifford, and P. Hutchings. 2019. Autonomy, Authenticity, Authorship and Intention in Computer Generated Art. In Proceedings of the 2019 on Creativity and Cognition (pp. 125-136). <https://doi.org/10.48550/arXiv.1903.02166>
- [12]. Elgammal, B. Liu, M. Elhoseiny, and M. Mazzone. 2017. CAN: Creative Adversarial Networks, Generating “Art” by Learning About Styles and Deviating from Style Norms. *arXiv preprint arXiv:1706.07068*.
- [13]. Caldarini, G., Jaf, S., & McGarry, K. (2021). A Literature Survey of Recent Advances in Chatbots. *Information*, 12(1), 41. <https://doi.org/10.48550/arXiv.2201.06657>
- [14]. Adamopoulou, E., & Moussiades, L. (2020). An overview of chatbot technology and its applications. *Machine Learning with Applications*, 2, 100006. <https://doi.org/10.1016/j.mlwa.2020.100006>
- [15]. Gill, S. S., Xu, M., Patros, P., Wu, H., Kaur, R., Kaur, K., Fuller, S., Singh, M., Arora, P., Parlikad, A. K., Stankovski, V., Abraham, A., Ghosh, S. K., Lutfiyya, H., Kanhere, S. S., Bahsoon, R., Rana, O., Dustdar, S., Sakellariou, R., Uhlig, S., & Buyya, R. (2024). Transformative effects of ChatGPT on modern education: Emerging era of AI chatbots. *Internet of Things and Cyber-Physical Systems*, 4, 19-23. <https://doi.org/10.1016/j.iotcps.2023.06.002>

- [16]. Costa, P. (2018). Conversing with personal digital assistants: on gender and artificial intelligence. *Journal of Science and Technology of the Arts*, 10(3), 2-59. <https://dx.doi.org/10.7559/citarj.v10i3.563>
- [17]. Lin, C.-C., Huang, A.Y.Q., & Yang, S.J.H. (2023). A review of AI-driven conversational chatbots implementation methodologies and challenges (1999–2022). *Sustainability*, 15(4012). <https://doi.org/10.3390/su15054012>
- [18]. Radziwill, N. M., & Benton, M. C. (2017). Evaluating quality of chatbots and intelligent conversational agents. *Journal of Quality Management*, 45, 34-49. <https://doi.org/10.1016/j.jqm.2017.05.001>
- [19]. Haase, J., & Hanel, P. H. P. (2023). Artificial muses: Generative artificial intelligence chatbots have risen to human-level creativity. *Empirical Studies of the Arts*. <https://doi.org/10.48550/arXiv.2303.12003>
- [20]. Nuruzzaman, M., & Hussain, O. K. (2018). A survey on chatbot implementation in customer service industry through deep neural networks. *Journal of Engineering and Applied Sciences*, 13(2), 45-51. <https://doi.org/10.3923/jeasci.2018.45.51>
- [21]. Urban, M., Dechterenko, F., Lukavský, J., Hrabalová, V., Svacha, F., Brom, C., and Urban, K. 2023. ChatGPT improves creative problem-solving performance in university students: an experimental study. *Computers and Education*, 215, 105031. <https://doi.org/10.1016/j.compedu.2024.105031>
- [22]. N. Davis et al. 2020. Empirically Studying Participatory Sense-Making in Abstract Drawing with a Co-Creative Cognitive Agent. *Proceedings of the ACM on Creativity and Cognition*. <https://doi.org/10.1145/2856767.2856795>
- [23]. C. Zhang et al. 2022. StoryDrawer: A Child–AI Collaborative Drawing System to Support Children’s Creative Visual Storytelling. *Frontiers in Psychology*, 13, 823078. <https://doi.org/10.1145/3491102.3501914>
- [24]. Qin, H. X., Jin, S., Gao, Z., Fan, M., & Hui, P. (2024). CharacterMeet: Supporting Creative Writers' Entire Story Character Construction Processes Through Conversation with LLM-Powered Chatbot Avatars. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24)*, May 11-16, 2024, Honolulu, HI, USA. <https://doi.org/10.1145/3613904.3642105>
- [25]. R. Hartati and E. B. Manullang. 2021. Implementation of Telegram Chatbot AI with Natural Language Processing (NLP) in Learning Creative Entrepreneurship to Develop Students' Creative and Innovative Competence. <https://doi.org/10.32734/lwsa.v7i2.2055>
- [26]. Ifelebuegu, A. O., Kulume, P., & Cherukut, P. (2023). Chatbots and AI in Education (AIED) Tools: The Good, the Bad, and the Ugly. *Education Sciences*, 13(2), 41. <https://doi.org/10.37074/jalt.2023.6.2.29>
- [27]. L. Clark. 2021. Towards "Creativity Amplification": or, AI for Writers, or Beating the System. *Proceedings of the ACM on Human Factors in Computing Systems*.
- [28]. Schmitt, O., & Buschek, D. (2021). CharacterChat: Supporting the Creation of Fictional Characters through Conversation and Progressive Manifestation with a Chatbot. In *Proceedings of the ACM Conference on Creativity and Cognition (C&C '21)*, June 22-23, 2021, Online. ACM. <https://doi.org/10.1145/3450741.3465253>
- [29]. Kantosalo et al. 2021. Empirically Studying Participatory Sense-Making in Abstract Drawing with a Co-Creative Cognitive Agent. *Proceedings of the 21st International Conference on Intelligent User Interfaces*. <https://doi.org/10.1145/2856767.2856795>

- [30]. Bickmore, T. W., & Picard, R. W. (2005). Establishing and Maintaining Long-Term Human-Computer Relationships. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 12(2), 293-327. <https://doi.org/10.1145/1067860.1067867>
- [31]. Ibarrola, F., Lawton, T., & Grace, K. (2022). A Collaborative, Interactive and Context-Aware Drawing Agent for Co-Creative Design. *Proceedings of the ACM on Human-Computer Interaction*. <https://doi.org/10.48550/arXiv.2209.12588>
- [32]. Hwang, A. H.-C., & Won, A. S. (2021). IdeaBot: Investigating Social Facilitation in Human-Machine Team Creativity. *Proceedings of the ACM on Human Factors in Computing Systems. Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. <https://doi.org/10.1145/3411764.3445270>
- [33]. Sun, Y., Li, X., Peng, J., & Gao, Z. (2021). Inspire Creativity with ORIBA: Transform Artists' Original Characters into Chatbots through Large Language Model. *Proceedings of the ACM on Creativity and Cognition*. <https://doi.org/10.48550/arXiv.2306.09776>
- [34]. Preuss, M., Liapis, A., & Togelius, J. (2014). Searching for good and diverse game levels. In *Proceedings of the IEEE Conference on Computational Intelligence and Games (CIG)*, Dortmund, Germany, 1-8. IEEE. <https://doi.org/10.1109/CIG.2014.6932908>
- [35]. Nicholas Davis. Human-Computer Co-Creativity: Blending Human and Computational Creativity. *Proceedings of the Ninth Artificial Intelligenceand Interactive Digital Entertainment Conference*
- [36]. Garcia, C. W. (2003). Drawing for the Absolute and Utter Beginner. Watson-Guptill Publications. ISBN: 9780823013951.
- [37]. Civardi, G. (2006). Drawing: A Complete Guide. Search Press. ISBN: 9781844485086.
- [38]. Cherry, E., & Latulipe, C. (2014). Quantifying the Creativity Support of Digital Tools through the Creativity Support Index. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 21(4), Article 21, 1-25. <https://doi.org/10.1145/2617577>
- [39]. Dewit, I., Latulipe, C., Dams, F., & Jacoby, A. (2020). Using the creativity support index to evaluate a product-service system design toolkit. *Journal of Design Research*, 18(5/6), 1-24. <https://doi.org/10.1504/JDR.2020.118670>.
- [40]. Holmes, S., Moorhead, A., Bond, R., Zheng, H., Coates, V., & McTear, M. (2019). Usability testing of a healthcare chatbot: Can we use conventional methods to assess conversational user interfaces? In *Proceedings of the 31st European Conference on Cognitive Ergonomics (ECCE 2019)*, 207-214. ACM. <https://doi.org/10.1145/3353562.3353577>
- [41]. Hennessey, B. A., Mueller, J. S., & Amabile, T. M. (2011). Consensual Assessment. In M. A. Runco & S. R. Pritzker (Eds.), *Encyclopedia of Creativity* (2nd ed., Vol. 1, pp. 253-260). Elsevier. <https://doi.org/10.1016/B978-0-12-375038-9.00044-1>

8. APPENDICES

8.1 APPENDIX 1. CONSENT FORM

Consent Form
Chatbot-Assisted Guidance in Digital Drawing.

This form seeks your consent to participate in a user study conducted by Leila Matayeva from University of Bremen. The purpose of this study is to test and compare different versions of the web drawing application with chatbot-assisted guidance.

As a participant, you will be asked to use a specific application developed for this research and complete a questionnaire that includes questions about your demographics (age, gender, experience) and other relevant information. Additionally, you will be asked to produce sketches within the application, which may be shown to other participants or individuals for further research purposes.

Your privacy and the confidentiality of your data are of utmost importance. Your personal information and any data you provide will be kept confidential and used solely for the purposes of this research.

Your participation in this study is entirely voluntary. You have the right to withdraw at any time without penalty.

Consent to Use of Data: By signing this form, you consent to the use of your data as described above. You understand that your sketches and anonymized demographic information may be used in research publications, presentations, or for educational purposes.

All data (demographics, questionnaire answers and sketches) will be published.

Signature:

I, _____, hereby consent to participate in the research study described above. I have read and understood the information provided and have had the opportunity to ask questions.

Participant's

Date: _____

Signature:

Researcher's

Date: _____

Signature:

8.2 APPENDIX 2. DATA PROTECTION FORM

**Data Protection Agreement
Chatbot-Assisted Guidance in Digital Drawing.**

This agreement outlines the measures taken to protect the personal and data privacy of participants in a user study conducted by Leila Matayeva from University of Bremen. The purpose of this study is to test and compare different versions of the web drawing application with chatbot-assisted guidance.

The researcher will collect personal data, demographics (age, gender, experience) and other relevant information, along with sketches produced within the research application.

The data collected will be used exclusively for the purposes of this research study. Personal identifiers will be removed from the data set to ensure anonymity before analysis.

Data will be retained for a period of 10 years after the completion of the study, after which it will be securely deleted.

Personal data will not be shared with third parties. Anonymized data and results may be shared for academic purposes, including publications and conferences.

Participants have the right to access their personal data, request correction of any inaccuracies, and withdraw their data from the study at any stage.

We will store and analyze the data for the Master thesis and for publications. Pseudonymised data will be published.

Agreement to Data Protection Measures:

I, _____, acknowledge the data protection measures outlined above and agree to the collection, use, and storage of my data as described.

Participant's _____ Signature: _____
Date: _____

Researcher's _____ Signature: _____
Date: _____

8.3 APPENDIX 3. DEMOGRAPHICS QUESTIONNAIRE FOR THE FIRST USER STUDY

Demographics

General information about the participants

* Indicates required question

1. How old are you? *

2. What is your gender? *

Mark only one oval.

- Female
- Male
- Non-binary
- Prefer not to say

3. How often do you create art such as drawing, painting etc.? *

Mark only one oval.

- Never
- Occasionally
- Sometimes
- Often
- Always

4. What exactly do you create (sketches, paintings, design, photography etc.) ? *

5. How often do you use AI tools (such as DALL-E, Midjourney, ChatGPT etc.)? *

Mark only one oval.

- Never
- Occasionally
- Sometimes
- Often
- Always

6. How often do you use AI tools for creative tasks (drawing, sketching etc.)? *

Mark only one oval.

- Never
- Occasionally
- Sometimes
- Often
- Always

7. How often do you create digital drawings? *

Mark only one oval.

- Never
- Occasionally
- Sometimes
- Often
- Always

8. Do you feel stuck or uninspired during creating art? *

Mark only one oval.

- Never
- Occasionally
- Sometimes
- Often
- Always

9. You consider yourself a creative person *

Mark only one oval.

- Strongly disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

10. You think that you lack some creativity in yourself *

Mark only one oval.

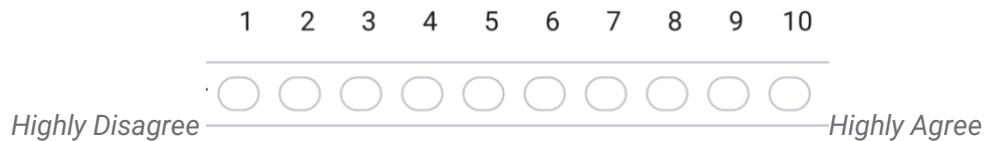
- Strongly disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

8.4 APPENDIX 4. FIRST USER STUDY QUESTIONNAIRE FOR EACH VERSION

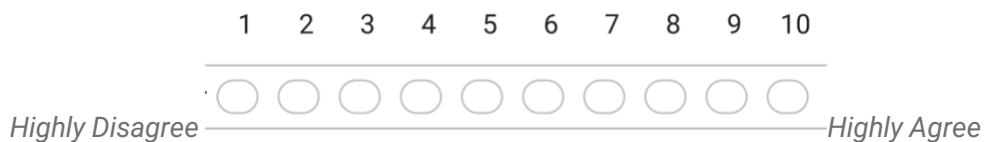
Questionnaire for All Chatbots Versions

1. What I was able to produce was worth the effort I had to exert to produce it.

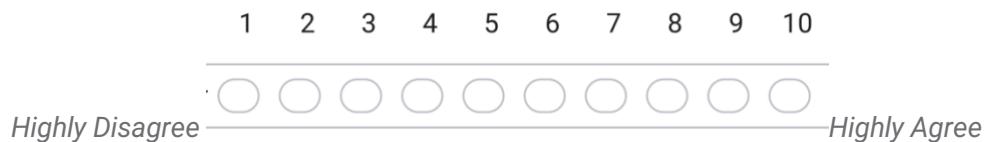
Mark only one oval.



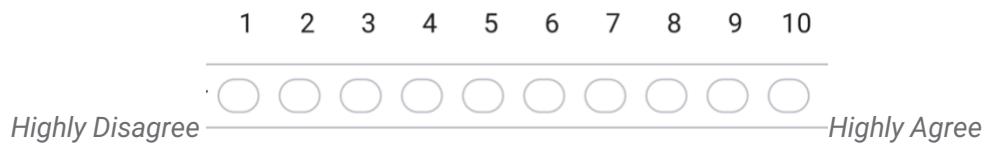
2. I would be happy to use this system or tool on a regular basis.



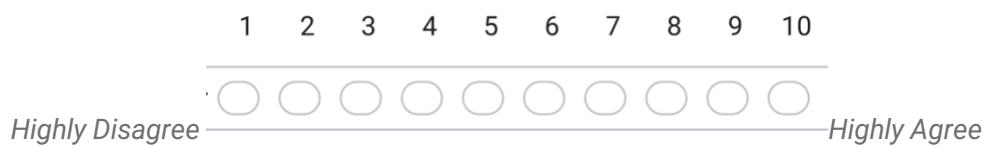
3. The system or tool allowed chatbot to work with me easily.



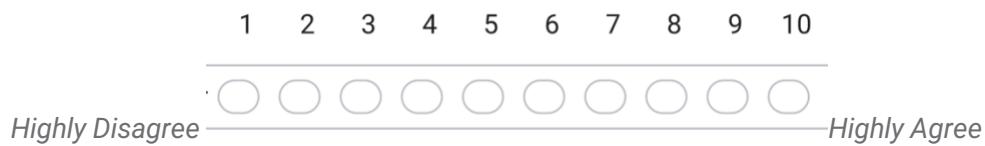
4. The system or tool was helpful in allowing me to track different ideas, outcomes, or possibilities.



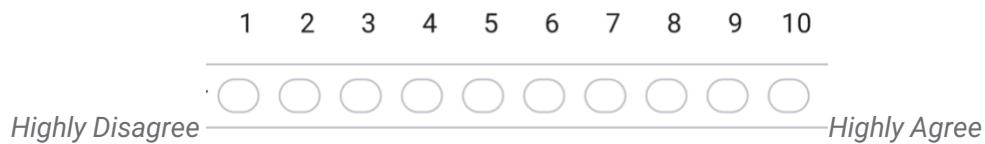
5. I enjoyed using the system or tool.



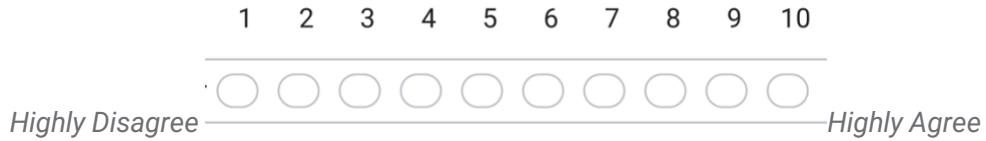
6. I was able to be very creative while doing the activity inside this system or tool.



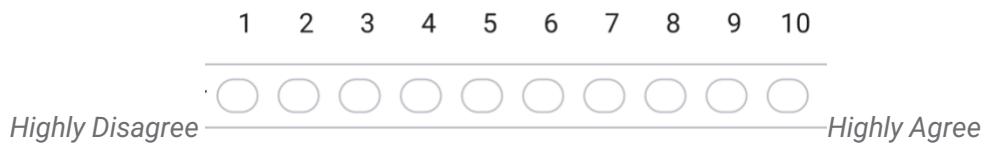
7. I became so absorbed in the activity that I forgot about the system or tool that I was using.



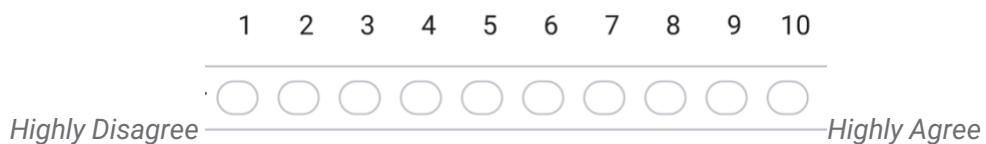
8. It was really easy to share ideas and thoughts with chatbot inside this system or tool.



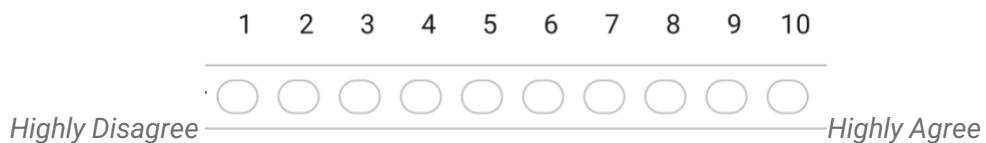
9. It was easy for me to explore many different ideas, options, designs, or outcomes, using this system or tool.



10. I was satisfied with what I got out of the system or tool.



11. My attention was fully tuned to the activity, and I forgot about the system or tool that I was using.



12. The system or tool allowed me to be very expressive.



CUQ

13. The chatbot's personality was realistic and engaging

Mark only one oval.

- Strongly disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

14. The chatbot seemed too robotic

Mark only one oval.

- Strongly disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

15. The chatbot was welcoming during initial setup

Mark only one oval.

- Strongly disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

16. The chatbot seemed very unfriendly

Mark only one oval.

- Strongly disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

17. The chatbot explained its scope and purpose well

Mark only one oval.

- Strongly disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

18. The chatbot gave no indication as to its purpose

Mark only one oval.

- Strongly disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

19. The chatbot was easy to navigate

Mark only one oval.

- Strongly disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

20. It would be easy to get confused when using the chatbot

Mark only one oval.

- Strongly disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

21. The chatbot understood me well

Mark only one oval.

- Strongly disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

22. The chatbot failed to recognise a lot of my inputs

Mark only one oval.

- Strongly disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

23. Chatbot responses were useful, appropriate and informative

Mark only one oval.

- Strongly disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

24. Chatbot responses were irrelevant

Mark only one oval.

- Strongly disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

25. The chatbot coped well with any errors or mistakes

Mark only one oval.

- Strongly disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

26. The chatbot seemed unable to handle any errors

Mark only one oval.

- Strongly disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

27. The chatbot was very easy to use

Mark only one oval.

- Strongly disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

28. The chatbot was very complex

Mark only one oval.

- Strongly disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

29. Chatbot inspired me to new ideas for drawings

Mark only one oval.

- Strongly disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

30. Chatbot messages kept me engaged to the drawing process

Mark only one oval.

- Strongly disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

31. Chatbot distracted me from drawing process

Mark only one oval.

- Strongly disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

32. Were there any specific features of the chatbot's behavior which made you feel inspired and creative? Why?

33. Were there any specific features of the chatbot's behavior which made you feel annoyed and distracted?

34. Did you feel stuck with no ideas/uninspired at any moment of the drawing process?

8.5 APPENDIX 5. END QUESTIONNAIRE FOR THE FIRST USER STUDY

End questionnaire

1. When doing this task, it's most important that I'm able to...

Mark only one oval.

- Be creative and expressive
- Become immersed in the activity

2. When doing this task, it's most important that I'm able to...

Mark only one oval.

- Enjoy using the system or tool
- Explore many different ideas, outcomes, or possibilities

3. When doing this task, it's most important that I'm able to...

Mark only one oval.

- Produce results that are worth the effort I put in
- Work with chatbot

4. When doing this task, it's most important that I'm able to...

Mark only one oval.

- Be creative and expressive
- Enjoy using the system or tool

5. When doing this task, it's most important that I'm able to...

Mark only one oval.

- Become immersed in the activity
- Produce results that are worth the effort I put in

6. When doing this task, it's most important that I'm able to...

Mark only one oval.

- Explore many different ideas, outcomes, or possibilities
- Work with chatbot

7. When doing this task, it's most important that I'm able to...

Mark only one oval.

- Be creative and expressive
- Explore many different ideas, outcomes, or possibilities

8. When doing this task, it's most important that I'm able to...

Mark only one oval.

- Enjoy using the system or tool
- Produce results that are worth the effort I put in

9. When doing this task, it's most important that I'm able to...

Mark only one oval.

Become immersed in the activity

Work with chatbot

10. When doing this task, it's most important that I'm able to...

Mark only one oval.

Be creative and expressive

Produce results that are worth the effort I put in

11. When doing this task, it's most important that I'm able to...

Mark only one oval.

Explore many different ideas, outcomes, or possibilities

Become immersed in the activity

12. When doing this task, it's most important that I'm able to...

Mark only one oval.

Enjoy using the system or tool

Work with chatbot

13. When doing this task, it's most important that I'm able to...

Mark only one oval.

Be creative and expressive

Work with chatbot

14. When doing this task, it's most important that I'm able to...

Mark only one oval.

- Produce results that are worth the effort I put in
- Explore many different ideas, outcomes, or possibilities

15. When doing this task, it's most important that I'm able to...

Mark only one oval.

- Become immersed in the activity
- Enjoy using the system or tool

16. It was easy to use drawing tablet for my sketching

Mark only one oval.

- Strongly disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

17. The interaction with the chatbot while using drawing tablet was smooth and convenient

Mark only one oval.

- Strongly disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

18. Which version of the drawing app you liked the most?

Mark only one oval.

- The one that was actively sending messages
- The one that was trying to guess what I'm drawing
- The one that was adding sketches to my drawing

19. Which version of the drawing app you liked the least?

Mark only one oval.

- The one that was actively sending messages
- The one that was trying to guess what I'm drawing
- The one that was adding sketches to my drawing

20. How do you find the chatbot interface for the digital drawing? Do you find it appropriate or you would prefer to have a different interface? Why?

8.6 APPENDIX 6. DEMOGRAPHICS QUESTIONNAIRE FOR THE SECOND USER STUDY

Consensual Assessment Technique Demographics

* Indicates required question

1. How old are you? *

2. What is your gender? *

Mark only one oval.

- Female
- Male
- Non-binary
- Prefer not to say

3. Do you have experience in working in an art related field? *

Mark only one oval.

- Yes
- No

4. Have you studied art related field? *

Mark only one oval.

- Yes
- No

5. How often do you create art such as drawing, painting etc.? *

Mark only one oval.

- Never
- Occasionally
- Sometimes
- Often
- Always

6. What is your educational background?

8.7 APPENDIX 7. CONSENSUAL ASSESSMENT TECHNIQUE QUESTIONNAIRE

Questionnaire for Consensual Assessment Technique

You will be given 3 art sketches by page. Please have a look at all 3 of them and rate how creative you think they are **compared to each other**.

DO NOT consider sketches from other pages, only 3 given to you on this page.

The definition of 'creativity' can be defined by you.

1. Compared to other two drawings, how creative do you think this drawing is? *



1 2 3 4 5 6 7

Very uncreative

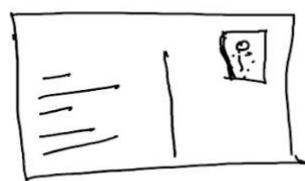
Very creative



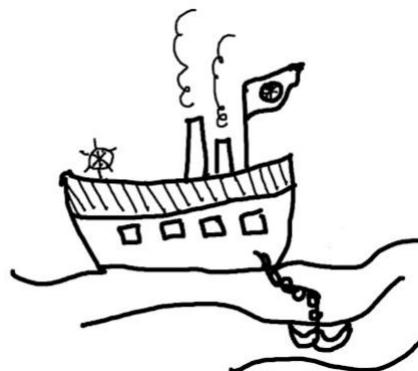
1 2 3 4 5 6 7

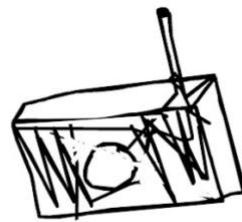
Very uncreative

Very creative



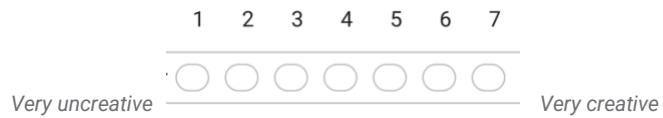
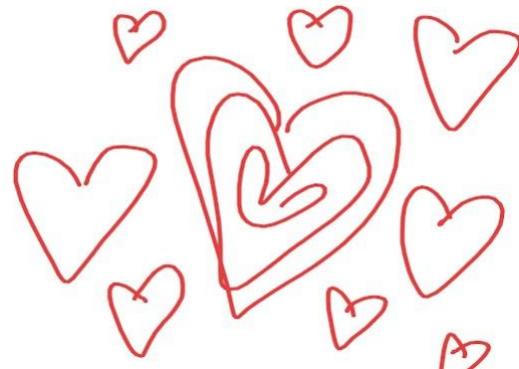
2. Compared to other two drawings, how creative do you think this drawing is? *



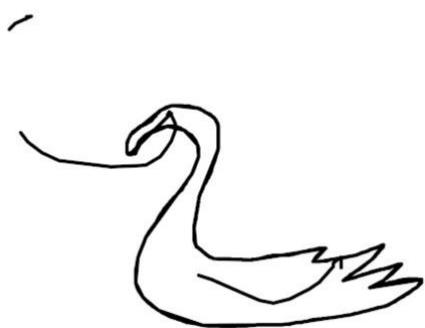
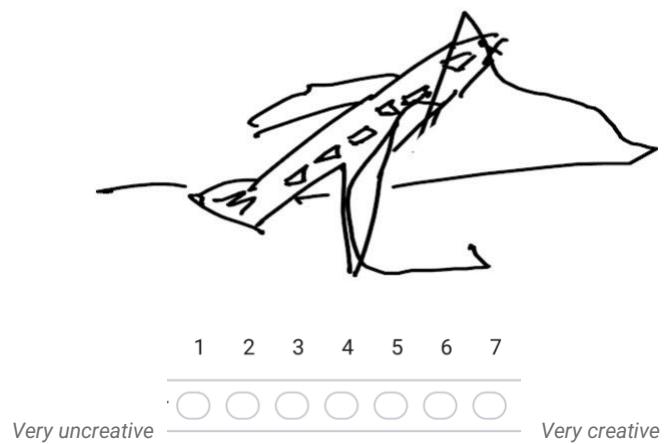
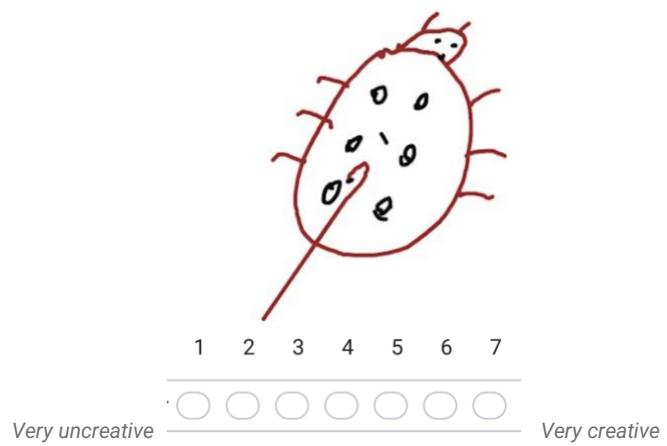


3. Compared to other two drawings, how creative do you think this drawing is? *





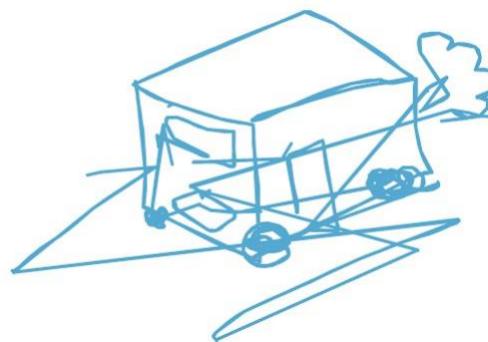
4. Compared to other two drawings, how creative do you think this drawing is? *





5. Compared to other two drawings, how creative do you think this drawing is? *

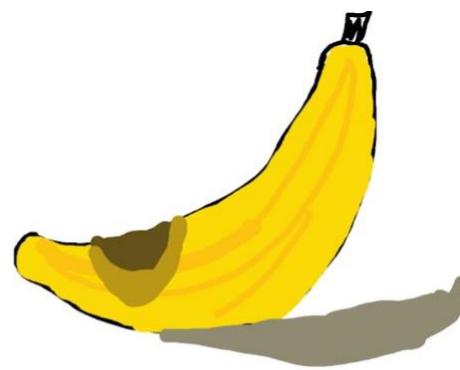




1 2 3 4 5 6 7

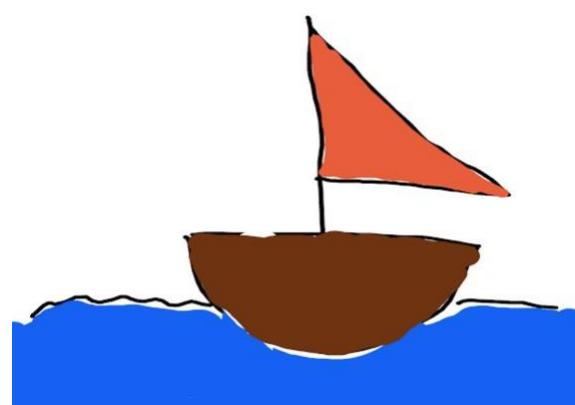
Very uncreative ————— Very creative

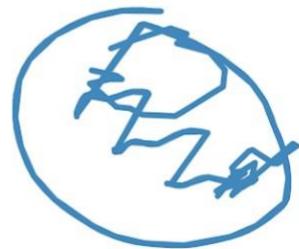
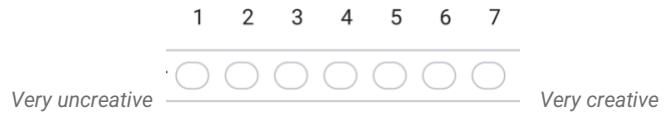
6. Compared to other two drawings, how creative do you think this drawing is? *



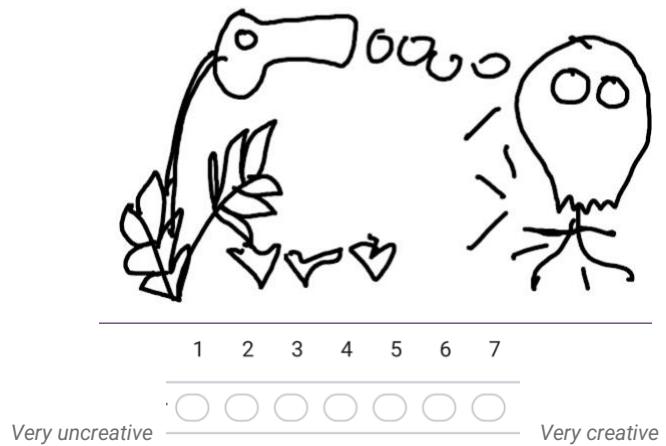
1 2 3 4 5 6 7

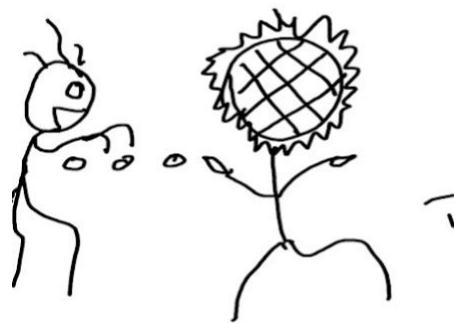
Very uncreative ————— Very creative





7. Compared to other two drawings, how creative do you think this drawing is? *

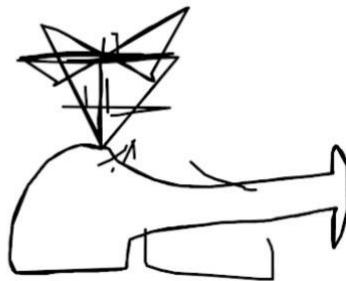




1 2 3 4 5 6 7

Very uncreative

Very creative



1 2 3 4 5 6 7

Very uncreative

Very creative

8. Compared to other two drawings, how creative do you think this drawing is? *



1 2 3 4 5 6 7

Very uncreative

Very creative



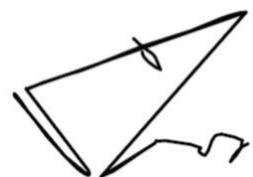
9. Compared to other two drawings, how creative do you think this drawing is? *





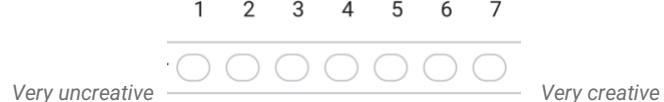
10. Compared to other two drawings, how creative do you think this drawing is? *





11. Compared to other two drawings, how creative do you think this drawing is? *





8.8 APPENDIX 8. DEMOGRAPHICS ANSWERS FOR THE FIRST STUDY

How old are you? Participant	What is your gender?	How often do you create art such as drawing, painting etc.?	What exactly do you create (sketches, paintings, design, photography etc.) ?	How often do you use AI tools (such as DALL-E, Midjourney, ChatGPT etc.)?	How often do you use AI tools for creative tasks (drawing, sketching etc.)?	How often do you create digital drawings?	Do you feel stuck or uninspired during creating art?	You consider yourself a creative person	You think that you lack some creativity in yourself	Versions' ordering
1	25	1	3	3	2	1	4	4	4	1
2	26	2	4	4	1	3	4	4	4	2
3	31	1	2	3	1	1	2	4	2	3
4	26	1	2	4	1	2	4	3	4	1
5	27	1	3	4	1	1	2	3	3	2
6	23	2	2	4	2	1	4	3	3	3
7	26	1	2	5	2	3	2	3	3	1
8	26	2	2	5	3	2	2	3	2	2
9	26	2	5	4	1	1	3	4	2	3
10	26	1	3	4	1	1	2	3	3	1
11	25	2	2	4	2	2	4	3	3	2
12	28	2	4	4	3	3	4	4	4	3

8.9 APPENDIX 9. ANSWERS TO THE ‘ACTIVE’ CHATBOT VERSION QUESTIONNAIRE (CSI, CUQ, ADDITIONAL QUESTIONS)

Participant number	What I was able to produce was worth the effort I had to exert to produce it.	The system or tool allowed chatbot to work with me easily.	The system or tool was helpful in allowing me to track different ideas, outcomes, or possibilities.	It was easy for me to explore many different ideas, options, designs, or outcomes, using this system or tool.		I was satisfied with what I got out of the system or tool that I was using.	My attention was fully tuned to the activity, and I forgot about the system or tool that I was using.
				It was really easy to share ideas and thoughts with the chatbot inside this system or tool.	I became so absorbed in the activity that I forgot about the system or tool that I was using.		
1	5	7	9	7	9	4	7
2	8	9	7	7	5	3	7
3	5	5	8	7	7	4	3
4	8	8	7	7	9	9	7
5	5	8	8	7	6	7	5
6	8	6	5	4	4	8	7
7	7	6	9	8	8	9	9
8	8	9	8	9	9	8	9
9	8	6	5	7	6	3	5
10	8	8	7	7	9	7	7
11	7	7	8	9	9	8	8
12	8	9	7	7	5	3	7

Participant number	The chatbot's personality was realistic and engaging	The chatbot was welcoming during initial setup	The chatbot seemed very unfriendly	The chatbot explained its scope and purpose well	The chatbot gave no indication as to its purpose	It would be easy to get confused by the chatbot	The chatbot failed to recognize a lot of my inputs well	Chatbot responses were useful, appropriate and informative	Chatbot responses were unable to handle any errors or mistakes	The chatbot coped well with any errors	The chatbot was very easy to use
1	4	2	4	2	3	3	4	1	3	2	3
2	4	2	4	2	4	2	4	2	5	4	2
3	4	3	4	2	4	2	4	3	4	2	3
4	4	3	4	1	4	3	5	2	4	2	3
5	4	3	3	2	3	4	4	3	4	3	4
6	3	5	5	1	4	3	2	3	2	4	2
7	4	4	4	2	4	4	4	2	4	2	3
8	4	2	4	2	4	2	4	2	4	2	2
9	4	2	4	2	4	2	4	3	3	4	2
10	3	3	4	2	3	2	3	2	4	3	2
11	4	2	5	2	4	2	5	2	3	4	1
12	4	3	3	2	3	4	4	3	4	3	2

Participant number	Chatbot inspired me to new ideas for drawings	Chatbot messages kept me engaged to the drawing process		Were there any specific features of the chatbot's behavior which made you feel inspired and creative? Why?	Did you feel stuck with no ideas/uninspired at any moment of the drawing process?
		Disagree	Agree		
1 Agree	Agree	Disagree	I like how it gave me steps of how to draw	It kept saying I am paused	No, it was neutral
2 Agree	Strongly Agree	Disagree	I like how it gave me steps of how to draw	maybe the chatbot interacted a bit too fast when there was no input for some time.	yes
3 Agree	Agree	Disagree	I like how it gave me steps of how to draw	I found the messages useful but they came up a little bit too often	Nope
4 Strongly Agree	Neutral	Neutral	I gave it a prompt for a u funny picture idea, and it did its job well	Too many messages	yes
5 Neutral	Agree	Disagree	I gave it a prompt for a u funny picture idea, and it did its job well	There were a lot messages and they were pretty much the same a lot of the times.	No
6 Agree	Neutral	Agree	Yes, adding shadows was a good idea	none	no
7 Neutral	Agree	Disagree	Good answers	no	no
8 Agree	Agree	Disagree	nice, it gave me clear advice	no	no
9 Agree	Agree	Disagree	I liked the step by step explanation of tips, without them being automatically applied onto my design. Like, I have the tips but it is up to my to decide if I want to use them	The paragraphs were too long at times.	Yes, the chatbot failed to give me specific tips.
10 Neutral	Disagree	Agree			
11 Agree	Agree	Disagree	it was helpful	too many messages sometimes	nit really
12 Agree	Neutral	Disagree			no

8.10 APPENDIX 10. ANSWERS TO THE ‘RECOGNIZING’ CHATBOT VERSION QUESTIONNAIRE (CSI, CUQ, ADDITIONAL QUESTIONS)

Participant number	What I was able to produce was worth the effort I had to exert to produce it.	The system or tool system or tool allowed chatbot to work with me easily.	The system or tool was helpful in allowing me to track different ideas, outcomes, or possibilities.	I was able to be very creative while doing the activity inside this system or tool.	It was really easy to share ideas and thoughts with chatbot inside this system or tool.	It was for me to explore many different ideas, options, designs, or outcomes, using this system or tool.	My attention was fully tuned to the activity, and I forgot about the system or tool that I was using.	
				I became so absorbed in the activity that I forgot about the system or tool that I was using.	I was satisfied with what I got out of the system or tool that I was using.	I was satisfied with what I got out of the system or tool that I was using.	I was satisfied with what I got out of the system or tool that I was using.	
1	7	8	9	9	9	9	6	8
2	8	7	6	9	5	4	10	7
3	6	5	4	3	3	4	2	5
4	8	8	6	8	10	10	7	7
5	6	7	6	5	5	8	4	4
6	8	7	10	9	10	6	10	8
7	9	9	8	8	8	9	8	9
8	4	6	8	3	8	7	4	7
9	9	9	8	8	10	8	8	7
10	8	8	7	9	8	9	7	7
11	5	5	8	7	7	4	3	6
12	6	5	8	8	7	4	4	6

Participant number	The chatbot's personality was realistic and engaging	The chatbot was welcoming during initial setup	The chatbot explained its scope and purpose well	It would be easy to get confuse	The chatbot gave no indication as to its purpose	The chatbot failed to understand me well	Chatbot response were useful, appropriate and informative	Chatbot failed to recognise a lot of my inputs	Chatbot responses were irrelevant	The chatbot coped well with any errors or mistakes	The chatbot seemed unable to handle any errors	The chatbot was very easy to use
1	4	2	4	1	3	2	4	2	4	2	2	4
2	4	2	4	2	4	2	5	1	2	5	3	4
3	2	4	4	2	4	3	2	4	2	2	3	4
4	3	4	4	1	3	3	4	4	2	4	2	2
5	4	2	4	3	4	4	4	3	4	3	3	3
6	5	3	5	1	4	3	5	1	4	4	5	1
7	4	4	4	2	4	2	4	3	4	2	3	2
8	2	4	4	2	3	2	4	3	2	4	3	4
9	4	2	4	2	3	2	4	2	4	2	4	2
10	4	3	3	2	3	4	4	3	4	3	4	2
11	2	1	5	5	1	2	5	1	1	5	3	4
12	3	4	5	1	4	2	5	2	4	3	3	4

Participant number	Chatbot inspired me to new ideas for drawings	Chatbot messages kept me engaged to the drawing process	Chatbot distracted me from drawing process	Were there any specific features of the chatbot's behavior which made you feel inspired and creative? Why?		Did you feel stuck with no ideas/uninspired at any moment of the drawing process?	Were there any specific features of the chatbot's behavior which made you feel annoyed and distracted?
				Agree	Disagree		
1 Agree				I like how the behaviour was friendly with emojis, and was interesting to see what the chatbot guesses what I am drawing.	It was guessing what I was drawing, but did not give me additional inputs that I can try to draw based on the object the chatbot thinks it is.	No, it gave me a starting point to draw something in the beginning	no
2 Agree				Disagree	no	it answered way to much. the responses were somehow irrelevant or wrong.	yes.
3 Disagree				Agree	Disagree	Not really, it was just that i expected more ai functionality but it was implemented in another version.	Nope
4 Agree				Agree	Disagree	In this version not really	Yes
5 Disagree				Agree	Agree	The chatbot suggested me to add color which I didn't add before. After that it suggested me to add shades which was a cool suggestion, too.	While it recognized correctly (which I told them) and I was just adding color, it continued guessing which I found confusing.
6 Neutral				Agree	Disagree		Yes a bit
7 Agree				Agree	Disagree		
8 Agree				Agree	Disagree	even though it recognized my ideas in different way, but giving me some interesting points which is great to make things creative	never recognized my painting no
9 Strongly Agree				Strongly Agree	Disagree	Yes, most of its guesses were correct. This gave me the stimulation to try new designs.	Not that I can remember.
10 Neutral				Neutral	Disagree		No, it was pretty smooth.
11 Agree				Neutral	Disagree	i wish it was guessing several objects	no
12 Agree				Disagree	Neutral	it was fun to use it	No

8.11 APPENDIX 11. ANSWERS TO THE ‘DRAWING’ CHATBOT VERSION QUESTIONNAIRE (CSI, CUQ, ADDITIONAL QUESTIONS)

Participant number	What I was able to produce was worth the effort I had to exert to produce it.	The system or tool allowed chatbot to work with me easily.	The system or tool was helpful in allowing me to track different ideas, outcomes, or possibilities.	I was able to be very creative while doing the activity inside this system or tool.	If I became so absorbed in the activity that I forgot about the system or tool that I was using.	It was really easy to share ideas and thoughts with chatbot inside this system or tool.	It was easy for me to explore many different ideas, options, designs, or outcomes, using this system or tool.	I was satisfied with what I got out of the system or tool.	My attention was fully tuned to the activity, and I forgot about the system or tool that I was using.
1	5	5	4	4	6	5	6	6	7
2	3	4	2	2	7	3	3	9	2
3	3	3	3	2	7	4	3	3	2
4	4	5	6	7	3	8	8	5	4
5	5	2	7	3	5	4	5	3	3
6	6	3	7	9	5	6	1	4	7
7	8	8	8	8	8	8	8	8	8
8	4	3	3	3	3	3	3	3	3
9	8	2	4	7	8	7	10	3	7
10	4	3	3	2	3	4	4	3	4
11	6	7	6	5	5	8	4	7	4
12	6	3	7	9	5	6	1	4	7

Participant number	The chatbot's personality was realistic and engaging	The chatbot was welcoming during initial setup	The chatbot explained its scope and purpose well	The chatbot gave no indication as to its purpose	It would be easy to get confused by the chatbot	The chatbot failed to recognize a lot of my inputs	Chatbot responses were useful, appropriate and informative	Chatbot responses were irrelevant or inappropriate	The chatbot coped well with any errors or mistakes	The chatbot seemed unable to handle any errors	The chatbot was very easy to use
1	3	3	4	2	3	2	4	4	3	2	3
2	4	2	5	2	4	2	5	2	3	2	4
3	2	4	4	2	4	2	4	3	3	2	4
4	3	4	5	1	4	2	5	2	4	3	3
5	4	2	3	2	4	4	4	2	5	2	4
6	4	4	5	1	5	3	2	4	2	2	4
7	5	2	4	2	4	4	4	4	2	4	3
8	2	4	4	2	4	3	2	2	4	2	2
9	4	2	4	1	4	3	2	4	1	5	2
10	3	5	4	2	4	2	4	4	3	5	2
11	2	4	5	1	5	3	4	4	2	4	3
12	3	2	2	2	4	2	3	3	2	4	2

Participant number	Chatbot inspired me to new ideas for drawings	Chatbot messages kept me engaged to the drawing process	Chatbot distracted me from drawing process	Were there any specific features of the chatbot's behavior which made you feel inspired and creative? Why?	Did you feel stuck with no ideas/uninspired at any moment of the drawing process?

8.12 APPENDIX 12. ANSWERS TO THE SECOND USER STUDY

Participant number	How old are you?	What is your gender?	Do you have experience in working in an art related field?	Have you studied art related field?	How often do you create art such as drawing, painting etc.?	What is your educational background?	active	recognizing	drawing
1	23	2	0	0	4	Bachelors in computer science	5.181818182	3.818181818	2.636363636
2	30	1	0	0	1	Master in Chemistry	5.727272727	5	1.818181818
3	24	2	0	0	2	Master's degree secondary school	4.181818182	3.181818182	4.909090909
4	24	2	1	1	3	diploma/realschulabschluss	5.181818182	4.727272727	4.090909091
5	26	2	0	0	3	Software Engineering	4.181818182	3.272727273	2.727272727
6	24	2	0	0	2		4.181818182	3.818181818	4.545454545
7	28	2	1	0	2	Master of Arts	3.818181818	2.818181818	2.545454545
8	26	2	1	1	1	International relations + art management	4	3.545454545	4.090909091
9	25	3	1	1	5	Design	5.363636364	4.272727273	4.545454545
10	24	2	0	0	2	Bachelor's and Master's in logistics	4.818181818	4.181818182	1.636363636
11	24	2	0	0	2		3.545454545	3.818181818	3.454545455
12	17	2	1	0	3	High school	3.545454545	2.818181818	3
13	29	2	1	1	4	MA and MS Architecture	4.454545455	3.363636364	3.636363636
14	27	2	1	0	3	Computer Science	4.090909091	4.090909091	4.727272727
15	26	1	0	0	2	Neuroscience	4.727272727	4.454545455	4.181818182
16	30	1	1	1	1	Masters of science	4.090909091	3.272727273	1.636363636
17	25	2	0	0	3	Natural science in Bachelor's and Master's	3.636363636	2.454545455	3.636363636