



Chatbots and Art Critique: A Comparative Study of Chatbot and Human Experts in Traditional Chinese Painting Education

Jinyu Liu
Durham University, UK
vvmj34@durham.ac.uk

Effie Lai-Chong Law
Durham University, UK
lai-chong.law@durham.ac.uk

Hubert P. H. Shum
Durham University, UK
hubert.shum@durham.ac.uk



Figure 1: “The Four Gentleman”, plum, orchid, bamboo, and chrysanthemum in Traditional Chinese Painting

ABSTRACT

Driven by the recent incorporation of chatbots into art education, art critique as a key factor in this realm poses distinct challenges and opportunities for this technology intervention. This study investigates the efficacy of chatbot-generated critiques in art education and compares them to critiques from human specialists. We conducted an empirical analysis by evaluating both chatbot and human critiques of traditional Chinese paintings, focusing on “The Four Gentlemen”, plum, orchid, bamboo, and chrysanthemum, representative themes in traditional Chinese art (Figure 1). Traditional Chinese paintings provide an ideal subject for this research owing to their metaphorical meanings and cultural depth, which provides a complex context for evaluating the nuanced capabilities of chatbots and human critiques. Our analysis looked into several critical dimensions of critique quality. The preliminary results suggest that while chatbot critiques are detailed and relevant, they lack the depth and emotional engagement that are characteristic in human critiques. Human critiques appeared to be way more nuanced and contextual with more emphasis on the symbolic and cultural layers of the artwork. The findings of this research signify that chatbots can play a role in art education and elevate the learning experience through immediate and structured critique. However, they are not able to replace the deeper insights offered from human critiques.

Accordingly, this study recommends a hybrid approach: chatbots act as a supplementary pedagogical role, enriching art education and critique. The conclusions suggest further research into the integration of chatbots within art educational settings is needed, encouraging refinement of the design and application of chatbots in creative domains.

CCS CONCEPTS

• Human-centered Computing; • Human Computer Interaction (HCI);

KEYWORDS

Chatbots, Art Education, Painting Education, Art Critique, Large Language Model

ACM Reference Format:

Jinyu Liu, Effie Lai-Chong Law, and Hubert P. H. Shum. 2024. Chatbots and Art Critique: A Comparative Study of Chatbot and Human Experts in Traditional Chinese Painting Education. In *Nordic Conference on Human-Computer Interaction (NordiCHI 2024)*, October 13–16, 2024, Uppsala, Sweden. ACM, New York, NY, USA, 13 pages. <https://doi.org/10.1145/3679318.3685385>

1 INTRODUCTION

The integration of chatbots within educational spheres marks a great paradigm shift in teaching methodologies. In the field of education, chatbots democratise access to knowledge by providing immediate, scalable feedback [1]. As their capabilities continue to grow with advances in natural language processing and machine learning, new opportunities and challenges arise in pedagogical practices. However, the application of chatbots to the field of art



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

NordiCHI 2024, October 13–16, 2024, Uppsala, Sweden

© 2024 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-0966-1/24/10

<https://doi.org/10.1145/3679318.3685385>

critique remains largely unexplored. In this context, the study has two open research questions:

- How can the effectiveness of the art critiques generated by a chatbot be assessed in the context of traditional Chinese painting?
- Do the art critiques developed by a chatbot reach the level of depth and context understanding that a human expert can achieve?

Technological advances have provided chatbots with the ability to conduct detailed analyses of text and imagery [1], potentially reshaping how students interact with art critique. Despite their technological capabilities, it is an open question whether chatbots can provide the empathetic and delicate critique necessary for art education. This study aims to fill this gap by comparing the critiques from chatbots and human critics on traditional Chinese artworks, specifically 'The Four Gentlemen' including plum, orchid, bamboo, and chrysanthemum. These topics are selected not only for their important role in Chinese art but also for their profound symbolic meanings which reverberate across cultures [30], providing a strong base to test the capability of chatbots to handle complex artistic and cultural subjects. Meanwhile, traditional Chinese painting holds an important position not only within China but serves as an essential bridge in cultural art education across the globe. At present, art education in China has experienced a boom and the market is open to the incorporation of new technologies [37]. Many students are interested in learning and creating traditional Chinese art but usually have no access to qualified and professional instructors [38]. Therefore, chatbots grant an unprecedented opportunity to fill this existing educational gap through offering instant critique and guiding the learning process of a broad audience.

This is an overview of the paper: We first review the current development of chatbots with a focus on the critique function in the domain of art education. It is followed by the methodology section that provides the procedure for conducting a comparative analysis between the critiques of chatbots and human-provided critiques on traditional Chinese artwork. The results on the effectiveness of chatbots in providing art critiques are reported next. Finally, the broader implications of these findings are discussed. Three main contributions are made in the paper:

- We conducted an empirical comparison of chatbot and human-generated art critiques, obtaining specific strengths and weaknesses of each category. Unlike existing studies that treat chatbot applications for education broadly, this study addresses the specialised area of painting within art education.
- We introduced a structured six-dimensional framework for evaluating art critiques that could be generalised both in future research and practical applications in art education. This framework contributes to a more systematic way of judging critique quality along the different dimensions.
- We focused on traditional Chinese painting to test the capability of chatbots to handle culturally complex art forms. We also identified the potential as well as limitations in understanding and interpreting cultural aspects through chatbots.

2 BACKGROUND AND RELATED WORK

2.1 The Emergence of Chatbots in Art Education

Prior to the recent advancements in Natural Language Processing (NLP) and Machine Learning (ML), chatbots were simpler as they could not fully understand user input but now it is becoming increasingly possible [7]. Currently, chatbots are able to provide support on a wide scale in education such as providing course details [8], thereby acting as bridges between students and academic staff. In addition, they are integrated into e-learning platforms to provide instant assistance to learners and simplify the educational support process [9].

The artistic field has been dramatically influenced by chatbots with powerful abilities such as generating artistic images from text descriptions [10]. Recently, the application of chatbots in art has moved beyond creation into interpretation of art, as scholars examine the twofold capacities they can take up: from creation to critic [11]. Diversifying the purposes of chatbots in art education will be a future goal to enrich the learning experience [12].

We employed the ICAP framework (Interactive, Constructive, Active, and Passive) by Chi and Wylie to justify our theoretical basis for applying chatbots for art critiques. The ICAP framework classifies cognitive engagement into four modes: passive, active, constructive, and interactive [26]. Chatbots can facilitate engagement across all the following modes, making it an appropriate technology for art critiques:

Passive mode: Students receive and remember the learning content in their memory. In our case, chatbots are able to present information and critiques, for example, concerning technical aspects of an artwork that can be read by students.

Active mode: Students recognise the learning material without interpreting it. In this mode, chatbots will ask students to complete tasks related to representing elements of artwork or comparing their works with others.

Constructive mode: Students use their learning content to reflect or explain. In the realm of art critique, chatbots can prompt students to explain what they see in the artwork and have a dialogue around that.

Interactive mode: Students utilise the learning material with a partner. This can be achieved when chatbots facilitate a dialogue with multiple students. This dynamic interaction not only deepens students' understanding but also enhances their critical thinking skills to engage in sophisticated art critiques.

In painting education, chatbots have raised a lot of interest and turned a new page in creativity. Liu [13] indicates that their applications are not only able to create new artistry but can also simplify traditional techniques of painting, saving much time for the artists in doing the work. With the proliferation of related applications such as Midjourney or Dall-E, these tools are powerful for their potential to light up enduring creativity and arm students for the digital age [14]. Generative chatbots enhance creativity by allowing students to try out ideas quickly and expand their creativity [2]. Besides, they prepared students for working in a digital-centric professional world [2]. By understanding these abilities, educators can guide students to better learning results to meet the new expectations in the evolving art fields.

However, there are many challenges and debates about the role of chatbots in art education. Great care has to be taken to address issues of privacy and bias based on human-chatbot collaboration [15]. Chatterjee [16] reflects upon the changing relationship between traditional art and chatbots, stating that the ability of chatbots to facilitate artistic production requires a reassessment of the nature of art. He envisions a symbiotic future where human creativity and chatbots can augment each other to redefine our notions of beauty, creativity, and even art itself. Another major challenge would be to redefine the role of the art instructor in this new learning environment. Traditionally, teachers teach knowledge through direct interaction with students. Assuming the responsibility of certain roles being played by a chatbot, the role of a teacher might become more facilitative. This can also change the relationship between teacher and student. Careful consideration has to be given to retain human elements of mentorship [32]. Furthermore, current chatbots can still exhibit shortcomings and errors. Mistakes in image recognition or sentiment analysis could diminish the quality of feedback given to students [29]. The reliability and accuracy of chatbots are indispensable to safeguard credibility. Finally, as emphasised by the sociocultural theory about the role of social interaction and collaboration in learning, extensive use of chatbots will reduce the opportunities for students to interact and collaborate with peers [32], hence influencing students' social cognitive development in their art learning.

2.2 The Significance of Critique in Art Education

Critiques are integral parts in education since they foster the spirit of critical thinking and reflective learning [19]. Art critique presents an impetus for discussion, inviting students to look at their creative work critically and promoting the development of evaluative skills essential for artistic progression [17]. It underscores the melding of critical thinking, creation, and evaluation within art education [20].

The seminal work by Wolff and Geahigan in 1997 gives an overview of art critique in an educational setting [18]; it points to its role of enhancing analytical skills related to the history, production, and aesthetics of art. Art critique could range from practical considerations to theoretical concerns, ascertaining its essential place within art education [18]. Fitch continues this dialogue by demanding a 're-calibration' of critique quality to better serve educators and students [19]. The quality of critiques is closely linked with the effectiveness of pedagogical practices and determines achievement in education outcomes [20].

An art critique is an integration of distinct elements [18]. Based on synthesis from existing research and literature related to art education and creative processes, the following core principles would normally form up an effective art critique: Critique must be lucid and actionable; it should be accessible to artists at different skill levels [21]. Critiques should cover the salient components of the artwork including themes and compositional elements that have a strong influence on piece reception and execution [22]. To put an artist on her or his growth trajectory, critiques should do more than indicating what needs to be refined but rather offer constructive advice and encouragement [35]. An effective critique

should encompass technique, composition, colour dynamics, thematic nuance, and emotive conveyance, providing substantial depth for understanding an artwork [23]. Through specific examples and emphasising certain features of the work, the critique becomes more explicit and concrete, and easier to understand by the artists being critiqued [4].

In the realm of traditional Chinese painting, the subjects like "The Four Gentlemen" (plum blossoms, orchids, bamboos, chrysanthemums) bring rich symbolic meanings. From the stroke of the brush down to the use of colour, everything is associated with culture. Therefore, cultural depth in critique becomes more necessary [3]. Effective critiques in this field are more than technical assessment; they involve metaphorical layers bringing out aesthetic expressions to cultural spirit. This deep engagement with the artwork's cultural layers plays a crucial role in bringing out nuanced understanding and appreciation of traditional Chinese painting [3]. Therefore, critique of traditional Chinese painting should be culturally considerate and sensitive in ensuring that critique coheres to the profound symbolic meanings intrinsic in a work. This comprises examination at technique and composition levels, as well as thematic nuance [36]. They are both necessary for complete appreciation of the artist's work.

2.3 Chatbot and Human Critique Mechanisms in Education

The integration of chatbots into critique mechanisms has drawn great attention about their potential towards enhancing learning outcomes. Some are on a particular focus of personalised critique systems facilitated by chatbots to enhance academic performance and stimulate deeper exploration within the learning processes [24]. Other researchers like Tubino and Adachi [25], revisited the role of chatbots in developing critique literacy. They addressed the potential of chatbots to enhance students' understanding by the use of automated critique tools within a university setting.

Comparative analyses between chatbot-generated versus human-provided critiques have shown their different advantages and limitations about effectiveness, appropriateness, and user satisfaction. For example, Ndukwe et al. [27] examined automated grading systems that incorporate chatbots for interacting with students and rendering critique. The efficacy was then tested against traditional human grading. Fidan and Gencel [28] performed an empirical study on the efficiency of the chatbot systems concerning instructional settings, with a focus on critique timeliness and modalities. The results from this study suggest that in most of the digital learning environments, chatbots are capable of providing timely and effective critiques and increasing the engagement of learners. In efforts to determine the overall quality of critiques provided by chatbots, several frameworks have been documented, such as the Analytic Hierarchy Process, by Radziwill and Benton [31].

Chatbots' critique function has extended into art education, where they assist in providing critiques on artworks. Research has shown that chatbots can improve learning outcomes by providing timely and personalised feedback [1]. They can considerably aid in the art critique process by providing instantaneous feedback that immediately lets students apply feedback to their work. Unlike

human critics, who may be subjective in their arguments [5], chatbots can reply with consistency and objectivity due to predefined algorithms and databases.

Despite the above strengths for chatbots to provide critiques, researchers are still far from investigating the relative quality of chatbot- versus human-generated critiques within the art realm. This research will fill this gap by systemically evaluating chatbot-generated art critiques against those generated by humans on traditional Chinese paintings. This comparative analysis will help establish potential chatbot roles in art critique, guiding pedagogical strategies and technological integration that can be taken in future art domains.

3 METHODS

3.1 Material Selection and Design

3.1.1 Artwork Selection. The sample in this study was chosen to focus on Bird and Flower Painting, which is a significant theme in traditional Chinese painting. The theme mainly portrays different natural elements like flowers and trees with symbolic meanings. These elements usually embody cultural and philosophical values associated with artistic notions toward nature, representing thematic concepts like beauty and harmony [6]. This study specifically explores “The Four Gentlemen” including plum, orchid, bamboo, and chrysanthemum, which are the most representative subjects in this theme [30].

All artworks chosen are students’ work from a traditional Chinese painting course instructed by the first author, based on their embodiment of classical attributes of “the Four Gentlemen” in the theme of Bird and Flower Painting. Systematic selection for diversity and quality was applied from an initial large sample of student works. The pool consisted of 50 traditional Chinese paintings which students produced during the course. These artworks were assessed on their technical execution, adherence to principles of traditional Chinese painting, and successful portrayal of themes regarding the Four Gentlemen. Our research team then conducted preliminary screening, which resulted in 10 high-quality paintings, which most represent the selected theme. From these 10 paintings, four were finally selected to represent each subject of the Four Gentlemen. The selection was based upon consensus among our research team, where every painting represented the unique characteristic associated with the subject. Permissions to use these works were sought from students, ensuring ethical considerations and respect for intellectual property.

3.1.2 Critique Collection and Analysis. We utilised art critiques provided by two specialised chatbots in the realm of painting instruction and critique. For the two chatbots, we chose My Art Assistant: Fine Art Painting [33] as Chatbot Expert 1 (CE1) and Art Critique: AI-driven Art Analysis [34] as Chatbot Expert 2 (CE2), with both using the GPT-4 language model and applying customised functions to develop meaningful art critiques. The first chatbot has processed more than one thousand conversations in the ChatGPT community. According to the conversations generated, it is one of the top popular chatbots in the painting education area. This level of engagement reflects its adoption and usefulness in providing art critique. Its high conversation count mirrors a strong preference by

users. The second chatbot is also outstanding in the art critic field, having generated a conversation count of more than 1000, way surpassing similar chatbot applications. Its user rating of 4.4 out of 5 further proves its quality and user satisfaction. Such a rating is far higher than most in its category and thus evidence not just proficiency but also a high level of user acceptance. From these metrics, the two chatbots appear to be credible tools for providing art critique and have the potential to provide critiques in this research.

We concurrently recruited two professional human experts. The two experts are experienced in giving feedback to students and are skilled in traditional Chinese painting techniques. For the selection of human experts, we initially identified a group of professional painters within our network. These painters were those with great expertise in traditional Chinese painting. We contacted around ten painters, all of whom were experienced in the creation and criticism of traditional Chinese painting. With a focus on their specialised knowledge of traditional Chinese art and critique, we finally chose two experts to engage as Human Expert 1 (HE1) and Human Expert 2 (HE2): One male and one female. They were both instructors at universities specialising in traditional Chinese painting. One has a Ph.D. in Chinese Art, while another has an M.A. in Painting with a direction in traditional Chinese painting. They both have more than twenty years of painting experience and over ten years of teaching traditional Chinese painting. Their teaching experience ensures they can provide insightful and detailed critiques.

The chatbots and human experts were asked to comment on the selected artworks in a consistent structure. They were provided with a digital version of each painting. The digital versions guaranteed that the chatbots and the human experts both received the same high image quality for their critique. Besides, a uniform structured critique template was prepared for both. It guided the critique process and ensured consistency. The template was based on six dimensions mentioned before: Overall Evaluation [35], Clarity & Understandability [21], Relevance & Applicability [22], Constructiveness [35], Comprehensiveness [23], and Specificity [4]. The human experts and chatbots were instructed to use the given template to develop critiques for each painting. They then utilised the input data to produce critiques based on the structured format. Each critique responded to the six specified dimensions mentioned above, which provided detailed feedback about different aspects of the painting. Initial critiques from human experts and chatbots were reviewed by our research team about their depth and relevance according to the criteria of this study.

This structured approach ensures that each critique is comparable across chatbot-generated and human-provided critiques, allowing for a nuanced analysis of each medium’s strengths and weaknesses. In total, 16 critiques were collected: 8 critiques from human experts and 8 critiques from chatbots. Each critique was comprised of six dimensions based on the art critique rubrics mentioned above. The critiques collected from both chatbots and human experts create a rich dataset for analysis, representing diverse perspectives for every art piece. This critique collection process ensures balanced comparison, placing emphasis on the different approaches and insights chatbots and human experts offer.

Table 1: Art Critique Template

Dimension	Aspect 1	Aspect 2
Overall Evaluation (OE)	Does it provide a concise summary of the artwork’s overall impact, noting its strengths? [35]	Does the critique identify its potential areas for enhancement and efficacy in communicating the artist’s intended message or theme? [35]
Clarity & Understandability (CU)	Is the artwork’s message or intent communicated with clarity? [21]	Are the themes or concepts presented accessible to the audience? [21]
Relevance & Applicability (RA)	How does the artwork align with its stated theme or underlying concept? [22]	Do the artistic elements utilised (e.g., colour, form, composition) serve to reinforce the artwork’s message? [22]
Constructiveness (CT)	Does the critique identify the artwork’s strengths? [35]	Does the critique offer precise recommendations for the artwork’s enhancement? [35]
Comprehensiveness (CH)	Does the critique examine the technical proficiency displayed in the artwork and assess the skillful application of techniques? [23]	Does the critique evaluate the composition? How do the elements and principles of design contribute to the artwork’s overall impact? [23]
Specificity (SF)	Does the critique highlight specific aspects of the artwork that are particularly effective or lacking? [4]	Does the critique provide in-depth commentary on these specific features? [4]

3.2 Design Process

3.2.1 Experimental Design. A balanced and impactful art critique should be based on a structured set of criteria that guides analysis. Here are key criteria used to judge the effectiveness of art critiques:

Clarity & Understandability: The critique should be lucid and direct, ensuring the understandability and executability for all skill levels of artists [21]. This makes the critique accessible, helping the process of learning.

Relevance & Applicability: Critique must resonate with the specific themes and composition of the artwork [22]. This alignment is crucial as it relates directly to the artwork’s impact and its execution, enriching the artist’s understanding and appreciation.

Constructiveness: The criticism should point out what needs improvement and also be constructive and supportive to encourage artistic growth [35]. This helps the artists polish their skill while diversifying their creative expression.

Comprehensiveness: Critique shall discuss all aspects of the work, including technique, composition, colour dynamics, and emotional depth [23]. In this way, a comprehensive judgment about the artwork is presented, with deep insights useful for the development of the artist.

Specificity: The critique should be specific with elements in the artwork so that it becomes useful in practice and understandable [4]. This kind of precision will make it clear to artists where adjustments are needed.

These criteria provide a balanced framework that would not only aid in the assessment of art but also help artists provide effective critique. Based on the above-discussed key criteria, we designed a templated approach with a set of structured questions to elicit critiques. Table 1 summarises the six dimensions with two detailed aspects for evaluation in each dimension. Examples of chatbots and human expert critiques against specific dimensions are shown in Table 2.

3.3 Evaluation

3.3.1 Participants for evaluation. To effectively evaluate human experts’ critiques and chatbot-generated critiques, we invited individuals who are familiar with traditional Chinese painting and its cultural nuances. We targeted Chinese artists who had received systematic training in art and were knowledgeable about Chinese cultural contexts. These requirements are critical for evaluating the collected critiques. We contacted 20 individuals in our network who were either artists or art educators with experience in providing art critiques. We told potential participants that they would participate in this study with their expertise in art critique practices. The following criteria for selecting participants include: 1) The evaluators should at least possess a bachelor’s degree in art; 2) The evaluators need to have experience in giving art critiques. After the selection, the final evaluation group consists of five participants: three females and two males, all holding at least a bachelor’s degree in art from well-known institutions. Their professional backgrounds vary, with two being art teachers, two freelance artists, and one in the gallery curatorial role. Such a diversity of the evaluation group assures that it brings to this study rich practical experiences. It is important to note that this was an independent study; no company sponsorship was involved to avoid conflict of interest. The selection process was performed ethically, with all participants agreed to participate and contribute to this research under informed consent. Table 3 shows the details of each participant’s educational background and current career in the art field. This table is provided for reference about the experience and qualifications of the evaluators in this study.

3.3.2 Preparation for Analysis. In our study, chatbot critiques were generated in English, while human expert critiques were originally in Chinese. We used English for chatbot-generated critiques because it enables us to keep our vision of designing chatbots with art critique functions that can help a global audience appreciate art.

Table 2: Examples of Critiques from Chatbots and Human Experts





Artwork Category	Chatbot Critique Example	Human Expert Critique Example
<p>Plums</p> 	<p>“The contrasts between blossoms and the simplicity of branches are helpful to focus attention on the subject. The flowers themselves have many useful details, like small dots to represent stamens, making them intricate.” (SF from CE1)</p>	<p>“Specific elements that stand out positively include the boldness of the branches and various opacities with the blossoms to create some visual interest. Subtle shifts from deep to bright red contribute a rhythm within flowers.” (SF from HE1)</p>
<p>Orchids</p> 	<p>“The technique demonstrated a skilled hand creating delicate balance between control and spontaneity. The composition follows traditional asymmetry and contributes to an impression of motion within the painting.” (CH from CE2)</p>	<p>“The technical execution is great with strokes that embrace both boldness and delicacy to the orchids. Through how the artist arranged plants and open space, he could achieve a sort of balanced and serene tableau in the composition.” (CH from HE2)</p>
<p>Bamboo</p> 	<p>“The greatest attribute of this work is its use of space. I would suggest considering how heavily the ink is laid onto the leaves. Adding a couple more detailed areas would add interest without breaking the minimalistic feel of the image.” (CT from CE2)</p>	<p>“One of the strengths of this artwork is the expression it achieves to show the spirit of the bamboo through the rhythm and direction of the brushwork. The feel of the brush suggests both the strength and resilience of the stalks and the grace of the leaves. For further improvement, what could be explored by the painter is a greater range of ink densities that will add a further dimension to the bamboo.” (CT from HE1)</p>
<p>Chrysanthemums</p> 	<p>“This painting foreshadows the lushness of a garden where chrysanthemums take center stage. The theme of natural beauty and the fleeting nature of life can be easily received by an audience with flowers used to signify such themes.” (RA from CE1)</p>	<p>“This painting depicts a bouquet of chrysanthemums intertwined with fluttering butterflies, which gives the scene liveliness and serenity. The impression conveyed in this painting is an elegant natural scene telling viewers about the transience of flora and fauna.” (RA from HE2)</p>

Table 3: Summary of Evaluators

No.	Academic Background	Professional Engagement
Evaluator 1 (E1)	Bachelor of Arts	Freelance Painting Artist
Evaluator 2 (E2)	Master of Fine Arts	University Art Instructor
Evaluator 3 (E3)	Bachelor of Fine Arts	Exhibition Curator
Evaluator 4 (E4)	Master of Art History	College Art Instructor
Evaluator 5 (E5)	Master of Fine Arts	Gallery Owner and Independent Artist

English is spoken everywhere around the world and is suitable for testing the functionalities that realise cross-cultural understanding of different art forms. This would give the chatbots’ users across different linguistic and cultural backgrounds. In terms of human-generated critiques, we tried to evoke the depth and professional

insight about traditional Chinese painting. Such complex appreciation often entails cultural perspectives that Chinese painters would be easy to express in Chinese. We recognised that requiring them to critique in English could compromise the quality of their cultural insight. Thus, we allowed the human experts to write critiques in Chinese to preserve the depth of their evaluations.

After collecting all the critiques, we had a carefully designed translation process that would preserve the integrity and subtlety of their original analysis. To ensure this critique remains intact across different languages, we had our translations done by professional translators with expertise in art to ensure accuracy. The back-translation method was important in providing a quality control step. Discrepancies could be detected through discussions among translators to reach consensus regarding the most faithful translations. Additionally, specialists in traditional Chinese art were consulted to ensure that translations were relevant culturally. Iterative reviews have polished these translations to truly reflect the essence of the original critiques. A specialised glossary of art terminologies was also established, which helps in ensuring coherence throughout the translation process. Moreover, to avoid any possible source bias that could occur during the analysis phase, we anonymised all critiques. This ensures that the origin of every critique remains unknown to formulate objective assessments.

4 RESULTS

4.1 Quantitative Data Analysis

4.1.1 Preparation for Quantitative Data Analysis. A numerical scale of 1 to 5 has been created for the purpose of quantitative analysis. Based on this scale, each critique will be scored against the criteria of effective critique defined above. Table 4 provides a comprehensive scoring rubric organised for readability and convenient reference for this study. This scoring rubric was provided to the five evaluators to score each critique.

4.1.2 Descriptive Statistics. Descriptive statistics were applied to the scores of chatbots and human critiques after the data collection. The results showed remarkable differences in the mean scores between chatbot and human critique on all six dimensions that were evaluated: Overall Evaluation, Clarity & Understandability, Relevance & Applicability, Constructiveness, Comprehensiveness, and Specificity. Human critiques consistently perform better than chatbot critiques, with higher mean scores in each dimension. For example, under Clarity & Understandability, the mean scores were 4.55 and 4.80 for HE1 and HE2, against 3.60 and 3.55 for CE1 and CE2. On the dimension of Specificity, HE1 and HE2 mean scores were 4.85 and 4.90 as opposed to 2.70 and 2.60 for CE1 and CE2. Among the chatbot critiques, the highest mean scores were obtained in Clarity & Understandability with 3.60 (CE1) and 3.55 (CE2), and Relevance & Applicability with 3.55 (CE1) and 3.40 (CE2). The implication is that chatbots are relatively more effective on these two dimensions compared with other dimensions. On the other hand, the lowest mean scores for chatbot critiques were obtained in Comprehensiveness, where the scores are 2.90 (CE1) and 2.70 (CE2), and Specificity with mean scores of 2.70 (CE1) and 2.60 (CE2); these are the dimensions where chatbots tend to be least effective.

The variability in scores, indicated by standard deviations (SDs), showed human critiques have less dispersion across most dimensions, suggesting more consistent evaluations. For example, on the Comprehensiveness dimension, the SDs for HE1 and HE2 were 0.44 and 0.50, compared with 0.55 and 0.47 for CE1 and CE2. Similarly, in Specificity, the SDs of human critiques were 0.37 (HE1) and 0.31 (HE2), far below the 0.47 (CE1) and 0.50 (CE2) observed for chatbot

critiques. Among the chatbot critiques, Clarity & Understandability (SD = 0.50 for CE1 and SD = 0.51 for CE2) and Relevance and Applicability (SD = 0.51 for CE1 and SD = 0.50 for CE2) exhibited relatively higher variability, indicating less consistency in these dimensions. On the other hand, Constructiveness (SD = 0.32 for CE1 and SD = 0.37 for CE2) and Overall Evaluation (SD = 0.49 for CE1 and SD = 0.47 for CE2) showed less variability, suggesting more consistent performance by chatbots in these dimensions. This illustrates that the critiques provided by human experts are not only more effective on average, but also more reliable. The higher variability in chatbot critiques suggests that while they can occasionally provide useful insights, their performance is less predictable and consistent compared to human reviewers.

Furthermore, for a more intensive understanding of the relative effectiveness of chatbot-generated versus human-provided critiques, we visualized the scoring data across different dimensions using bar plots with error bars (Figure 2). It effectively summarised the descriptive statistics and provided a clear graphical representation of the mean scores and variability. The blue bars indicate the mean score of human critiques, whereas the red ones represent that of chatbot critiques. Error bars reflect the standard deviation. Along all dimensions, the mean scores for human critiques are higher than those of chatbot critiques. The dimension in which chatbot critiques performed relatively better is Clarity & Understandability, while the least effective dimensions for chatbot critiques are Comprehensiveness and Specificity. Standard deviations of human critiques are lower indicating more consistent evaluations, whereas chatbot critiques have higher variability for some dimensions, especially Clarity & Understandability and Relevance & Applicability.

4.2 Qualitative Data Analysis

The first step of our qualitative analysis was grouping the critiques into two categories: those from human experts and those from chatbots. This was important for comparative analysis and required further development through the classification of the critiques by the artwork evaluated. It would help in a systematic extraction of data, with each critique aligning to its appropriate artwork and source category for subsequent in-depth analysis.

In our study, thematic analysis was conducted to enable the exploration of data from evaluator comments. Five evaluators participated in assessing the quality of critiques provided by both chatbots and human experts. These evaluators used the scoring rubric based on the six dimensions introduced in the experimental design section. After each evaluator scored every dimension for each critique, detailed comments for every dimension were also required. This provided a total of 480 comments: 240 comments of chatbot-generated critiques and another 240 comments of human-generated critiques. Finally, thematic analysis was conducted to derive common themes and insights related to the strengths and weaknesses of chatbots and human critiques.

This qualitative approach helped in identifying themes and patterns of assessment made by the evaluators. The process involved several distinct phases. First of all, all comments given by the evaluators were copied from different dataset files into one dataset. This was important in ensuring that no insights were missed, and

Table 4: Art Critique Scoring Rubrics

Dimension	1(Lowest)	2	3	4	5(Highest)
Overall Evaluation	The critique does not provide a clear overall impression of the artwork.	The critique gives a vague overall impression with significant areas lacking detail.	The critique offers an average overall impression, but some aspects are underexplored.	The critique provides a clear overall impression, with well-identified strengths and areas for improvement.	The critique presents a comprehensive overall impression that insightfully covers strengths, areas for improvement, and success in conveying the artwork's message.
Clarity & Understandability	The artwork's intention or message is not communicated clearly.	The message is somewhat communicated, but viewers may find it challenging to understand.	The artwork's message is generally clear with minor areas of confusion.	The artwork's intention is clear, and themes are understandable with little effort.	The artwork's message and themes are communicated with excellent clarity and are easily understood by viewers.
Relevance & Applicability	The critique does not relate its observations to the broader context of the artwork's themes, styles, or historical significance.	The critique makes some relevant observations, but often fails to connect them effectively to the artwork's context or art principles.	The critique is somewhat relevant and applies some art principles but lacks depth in making connections to the artwork's broader themes.	The critique offers relevant observations that are well connected to the artwork's themes and demonstrates a good understanding of art principles.	The critique provides highly relevant and insightful observations that are deeply connected to the artwork's context and excellently applies art principles to support its points.
Constructiveness	The critique lacks identification of strengths and does not provide constructive suggestions.	Some strengths are identified; suggestions for improvement are vague	The critique identifies strengths adequately and gives general suggestions for improvement.	The critique clearly identifies strengths and offers specific, actionable suggestions for improvement.	The critique excels in recognizing strengths and provides insightful, detailed suggestions for improvement.
Comprehensiveness	The critique does not discuss the technical execution or analyze the composition.	The critique mentions technical execution and composition but lacks detail.	The critique provides a fair discussion of technical execution and composition.	The critique offers a detailed analysis of the artwork's technical execution and composition.	The critique provides an exceptionally thorough analysis, covering all aspects of technical execution and composition.
Specificity	The critique fails to identify specific elements or provide detailed critique.	The critique identifies few specific elements and provides limited critique.	The critique identifies specific elements and gives moderate critique.	The critique specifically highlights standout elements and provides detailed critique.	The critique demonstrates exceptional attention to detail, thoroughly analyzing standout elements with comprehensive critique.

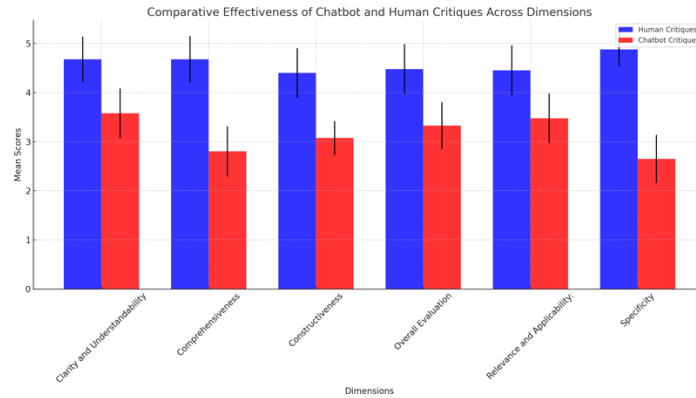


Figure 2: Comparison of Chatbots vs Human Critique Scores Across Dimensions

the analysis represented the critique from all evaluators. The initial coding process involved reading through each comment and looking for key phrases or concepts that emerged. This was a basic step toward identifying the descriptive and evaluative phrases which indicated focus areas or criteria used by the evaluators in their judgments. For example, comments such as “The critique provides a comprehensive view” and “The critique clearly conveys the themes of the artwork” reflected attention to comprehensiveness and clarity.

After the initial coding, focused coding was done where similar codes were grouped under categories representing broader themes. This step made it easier to organise the codes into more significant clusters reflecting overarching themes across all comments. Initial codes were finally grouped under the following three themes:

Effectiveness of Communication: This theme includes codes related to clarity and the ability to effectively communicate the main points of the critique (i.e. ability to convey the essence and symbolism of the artwork).

Analytical Depth: It includes codes demonstrating depth of analysis (i.e. insightfulness in tying visual elements to broader cultural themes).

Constructive Critique: It focuses on the utility of the critique (i.e. enhance the artist’s narrative or compositional approach).

Finally, these themes were synthesized to create an understanding of what criteria evaluators focus on in their judgments about the critiques. The synthesis not only helped point out what the evaluators focused on as their evaluative priorities, but also provided insights into those areas where chatbot-generated critiques need improvement.

4.2.1 Focused Coding and Thematic Analysis. To gain a deeper understanding of the criteria that evaluators focus on when evaluating critiques, we conducted a detailed analysis of the identified themes derived from focused coding.

Following the organisational stage, we took detailed coding against predefined themes established from our preliminary thematic exploration. These themes provided the framework for coding across six evaluative dimensions in critiques. Each of the critiques was examined and coded against these dimensions. Particular attention was paid towards how well each critique communicated

the artwork’s themes and intentions, the depth of insight into the artwork’s technical and thematic aspects, and the usefulness and balance of the critique (Table 5).

4.2.2 Comparative Analysis with Evaluator Comments. We performed a qualitative comparison of human and chatbot critiques for four different traditional Chinese paintings: Plums, Orchids, Bamboo, and Chrysanthemum. This analysis covers three relevant dimensions (Effectiveness of Communication, Depth of Analysis, and Constructive Critique) of critique quality that provide insights into strengths and limitations of each source (Table 6).

The analyses clearly show that human experts have been able to produce art critiques with deeper communication, and more profound and emotional analysis compared to chatbot critiques. While chatbot critiques perform well at providing clear and elaborate factual analysis, they have much less effective cultural or symbolic interpretation and emotional engagement. These findings show the critical areas where chatbot critiques need to be fine-tuned.

5 DISCUSSION

5.1 Revisiting the research questions.

This study focused on chatbot applications in art education, particularly in art critique. It explored innovative methods for assessing the effectiveness of art critiques generated by chatbots, focusing specifically on traditional Chinese painting. We tried to present standards to evaluate the depth of critiques based on an empirical analysis of art critiques. The results from the comparative analysis indicate that though chatbots were able to provide relatively reliable critiques, the content of the critiques often lacked interpretable depth and emotional engagement.

Quantitative and qualitative analyses provided deeper insights into the role and effectiveness of chatbots in art critique. Based on the analysis, the evaluators appreciated the clarity and factual analyses from chatbot feedback. Chatbot critiques have successfully identified detailed areas for improvement. This was quite useful for learning and refining technical aspects related to brushwork, colour usage, and composition. However, chatbots can easily lack nuanced cultural contexts that a human critique can provide in many cases. Evaluators also commented that feedback given by

Table 5: Description of themes and subthemes identified in the evaluators' comments.

Themes (Subthemes)	Description	Sample Quotes
Effectiveness of Communication	How effectively critiques convey the artwork's essence, themes, and symbolism	"It demonstrates the ability to convey the artwork's essence and symbolism." (E2)
Clarity	Captures comments that emphasise the clear presentation of the artwork's intentions and thematic elements	"It communicates the essence and symbolism of the artwork effectively," (E5) "The artwork's themes and intentions are articulated clearly." (E2)
Engagement	Highlights the critique's ability to engage the reader and vividly portray the artwork's message	"The critique adeptly captures the artwork's visual and thematic elements," (E3) "Effectively communicates the artwork's visual narrative to the audience." (E1)
Analytical Depth	Focuses on the depth of analysis regarding both the technical aspects and the thematic depth of the artworks	"It provides insightfulness in tying visual elements to broader cultural themes." (E2)
Technical proficiency	Discuss the critique's detailed analysis of the artwork's technical execution, including aspects like brushwork and colour blending	"The technical execution displays proficient brushwork and color blending," (E1) "The critique thoroughly discusses the technical execution." (E2)
Cultural and thematic insights	Reflects the critique's ability to deeply explore cultural or thematic layers of the artwork	"Deeply rooted in cultural symbolism, the critique unpacks the artwork's layers," (E5) "Astutely ties the artwork's visual elements to broader cultural themes." (E3)
Constructive Critique	The usefulness of the critique in providing actionable critique or enhancing understanding of the artwork	"The critique enhances the artist's narrative or compositional approach." (E3)
Actionable suggestions	Comments that highlight practical suggestions made by the critique to improve or deepen understanding of the artwork	"Provides useful suggestions for adding depth to the artistic narrative," (E2) "Constructive in identifying the artwork's compositional strengths." (E4)
Balanced criticism	Discuss how critiques balance praise with constructive critique to offer a well-rounded evaluation	"While acknowledging the artwork's strengths, the critique offers constructive ways to enhance the narrative," (E4) "Recognizing the artwork's strengths in composition while providing critique on thematic depth." (E1)

chatbots was devoid of emotional engagement compared with that given by human experts.

Such findings underline the shortcomings of current chatbot technologies in art critique and further emphasise that more advancement would be needed to boost their interpretative abilities.

Such an outcome also infers that while chatbots are useful in augmenting the critique part of art education, they are still some distance away from adequately replicating the human expertise in providing art critique.

Table 6: Critique Quality Comparison Based on Themes

Artwork Category	Effectiveness of Communication	Analytical Depth	Constructive Critique
Plums	Human critiques are vivid with sensual descriptions. The cultural symbolism creates an exciting reading experience when bringing in real depictions of the theme. ***	Human critiques go into complex subjects of colour vibrancy and form, contextualising them within broader cultural and artistic trends. ***	Human critiques offer pointed suggestions that respect and enrich the thematic integrity of the artwork. ***
	While effective at displaying the basic thematic and compositional details, nuanced language that enhances emotive engagement lacks in chatbot critiques.	Chatbots provide structured, fact-based analyses which are more descriptive than interpretive.	Chatbot critiques tend to provide general critique that is usually technically correct but lacking in personalised insight.
Orchids	Human reviews are noted for the emotional resonance that is characterised by evocative language. It tunders an emotional appeal to the aesthetic placement of the orchids. ***	Human critiques focus on the cultural symbolism of orchids, such as beauty and grace. ***	Suggestions from human critics often try to focus on the artwork’s emotional effect. ***
	While clear and accurate, chatbot critiques lack the same emotional depth compared with human critiques.	The chatbots focus on more technical aspects of the artist’s style, like brush strokes and colour use, rather than deeper symbolic interpretations.	Chatbot critiques focus on enhancing the technical aspects and fail to highlight emotional or thematic substance.
Bamboo	Human critiques deal much with the symbolism of bamboo and how it represents resilience. This is connected with the underlying theme present in the artwork. ***	The human critique takes into consideration some historical and cultural significance surrounding bamboo. The critique also links the painting to philosophical concepts. ***	Human critique is nuanced, considering artistic intent and viewer reception. ***
	While being descriptive of visual aspects, chatbots are not as effective at conveying the deeper meaning related to the theme.	Chatbots limit their analysis to observable elements without broader contextual interpretation.	Chatbot criticism is limited to visible technical enhancements.
Chrysanthemum	Human critiques create narratives that touch emotionally into the symbolism of chrysanthemums, especially on themes of longevity and autumn. ***	For human critiques, the use of colours and composition is interpreted as a reflection of life and seasonal themes. ***	Human critiques include creative suggestions which do enhancement to thematic portrayal. ***
	Chatbots describe visual characteristics effectively but lack narrative depth.	Chatbots make very detailed but still superficial observations.	Chatbots’ focus on technical improvements such as saturation or detail enhancement.

5.2 Practical and Methodological Implications

Our study provides an in-depth empirical comparison of art critiques provided by chatbots and human experts. This area has remained substantially underexplored within existing literature. Although much of the past research investigated chatbot applications within general educational settings, very few have examined or considered their application in the art critique domain. This paper provides new empirical data by focusing on traditional Chinese paintings that outline the strengths and weaknesses of chatbots’ art critique. The focus on traditional Chinese painting provides an additional cultural layer about chatbots in art education. Compared to previous studies, this research evaluates how well chatbots can

understand and critique artworks influenced by cultural symbolism. This opens new studies regarding the capabilities and limitations of chatbots to the understanding of cultural nuances in art criticisms.

Different dimensions of art critique are introduced in this research in the form of a six-dimensional framework. The framework provides a relatively clear guideline for evaluating the critiques and can be adopted in future research to ensure consistency in the analysis. However, it also raises questions of potential influence on the traditional art critique process. Art critique often thrives on the freedom to discuss art in an open-ended manner. The imposition of a structured framework could potentially restrict this exploratory nature, limiting the scope of critique to predefined

categories and possibly decreasing the subjective elements that are vital to art criticism. This tension between structure and freedom represents a crucial methodological consideration, suggesting that while structured frameworks can enhance the comparability of critiques, they must be designed thoughtfully to preserve the richness of spontaneous art discussion. The findings of this study also demonstrate how essential it is to develop the capabilities of chatbots from superficial interactions to complex communication situations like providing art critique. Future development efforts should focus on sophisticating their analytical capabilities to interpret and respond to nuanced aspects of art critique. Based on the study results, current chatbots can only serve as complementary tools in art critique. The educator could be assisted by chatbots to provide instant and high-quality feedback to a large group of students. In places with a lack of experienced art teachers, chatbots could be an answer to quality critique. For the larger perspective on how chatbots are impacting art education, they have the potential to reach a global audience, transcending geographical barriers and promoting cross-cultural exchange through art education.

5.3 Limitations and Future Studies

Preliminary results from this study raise a compelling need for more refined methodologies in the future. This study had a relatively small sample size of human experts, which is not representative enough to create variability to capture diversity in the critique provided by educators working in different fields. Future research shall consider an increase in the sample size by including diversified human experts. This would help to understand the overall differences between chatbot-generated critiques and those given by the art experts, hence providing a broader assessment of the chatbot's effectiveness.

Another limitation is the translation of the critiques and comments from Chinese to English. Although the professional translators ensured the accuracy, some aspects of the Chinese language, particularly concerned with cultural nuances, are hard to translate. This may affect the contextual richness of the critiques to be presented in English. In the future, experts and evaluators who all speak English could be used, so as not to affect the integrity of language and provide more accurate evaluation.

Future studies should diversify the range of art styles examined, going beyond paintings to sculpture, photography and even moving images. This would test the adaptability of chatbot critiques across different artistic expressions and help to determine whether their effectiveness generalises across different forms of art. Moreover, the rapid technical evolution of chatbots would provide new dimensions to evaluate their capacity for art critique.

6 CONCLUSION

This research provides valuable insights into both the capability and limitation of chatbots in providing art critique. While current chatbots demonstrated proficiency in handling straightforward descriptive tasks, they fell short in delivering the depth required for more nuanced art critique. The present research sets the scene for future studies in improving the interpretive abilities of chatbots and opens the path toward more sophisticated chatbot applications

in art education. We made a comprehensive analysis of the differences in effectiveness between the chatbot-generated and human-provided critiques on various dimensions of evaluation. The results from the data analysis showed an overall preference towards human critique, scoring higher than their chatbot counterparts across several dimensions.

REFERENCES

- [1] Eleni Adamopoulou and Lefteris Moussiades. 2020. An Overview of Chatbot Technology. In *Artificial Intelligence Applications and Innovations*, Springer, Cham. https://doi.org/10.1007/978-3-030-49186-4_31
- [2] Jiahong Su and Weipeng Yang. 2023. Unlocking the Power of ChatGPT: A Framework for Applying Generative AI in Education. *ECNU Review of Education* 6, (August 2023), 355–366. <https://doi.org/10.1177/20965311231168423>
- [3] Sherman E. Lee and Arthur Waley. 1959. An Introduction to the Study of Chinese Painting. *Journal of Aesthetics and Art Criticism* 17, 3 (March 1959), 390. <https://doi.org/10.2307/427817>
- [4] Matthew Bowman. 2021. Art criticism in the contracted field. *Journal of Aesthetics and Art Criticism* 79, 2: 200–212. <https://doi.org/10.1093/jaac/kpab003>
- [5] Eleni Gemtou. 2010. Subjectivity in Art History and Art Criticism. *Rupkatha Journal on Interdisciplinary Studies in Humanities* 2, 1 (March 2010). <https://doi.org/10.21659/rupkatha.v2n1.02>
- [6] Feifei Fan and Safial Aqbar Zakaria. 2024. The philosophy of color evolution in Chinese bird and flower paintings and its integration with new Chinese interior design. *Int. J. Philos. Cult. Axiol.*, 21, 1, 152–171.
- [7] Md.AI-Amin, Mohammad Shazed Ali, Abdus Salam, Arif Khan, Ashraf Ali, Ahsan Ullah, Md Nur Alam and Shamsul Kabir Chowdhury. 2024. History of generative Artificial Intelligence (AI) chatbots: past, present, and future development. *ArXiv: 2402.05122*, Retrieved from <https://doi.org/10.48550/arXiv.2402.05122>
- [8] Mendoza Sonia, Luis Martín Sánchez-Adame, José Fidel Urquiza-Yllescas, Beatriz A. González-Beltrán, and Dominique Decouchant. 2022. A Model to Develop Chatbots for Assisting the Teaching and Learning Process. *Sensors (Basel, Switzerland)*, 22. <https://doi.org/10.3390/s22155532>.
- [9] Saqib Hussain, Safa Hamood Al-Hashmi, Mazhar Hussain Malik and Syed Imran Ali Kazmi. 2023. Chatbot in E-learning. *SHS Web of Conferences*. <https://doi.org/10.1051/shsconf/202315601002>.
- [10] Joern Ploennigs and Markus Berger. 2022. AI art in architecture. *AI in Civil Engineering*, 2. <https://doi.org/10.1007/s43503-023-00018-y>.
- [11] Eva Cetinic and James She. 2021. Understanding and Creating Art with AI: Review and Outlook. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, 18, 1–22. <https://doi.org/10.1145/3475799>.
- [12] Fanwen Kong. 2020. Application of Artificial Intelligence in Modern Art Teaching. *Int. J. Emerg. Technol. Learn.*, 15, 238–251. <https://doi.org/10.3991/ijet.v15i13.15351>.
- [13] Mingyang Liu. 2023. Overview of Artificial Intelligence Painting Development and Some Related Model Application. *SHS Web of Conferences*. <https://doi.org/10.1051/shsconf/202316701004>.
- [14] Ahmad Faisal Choiril Anam Fathoni. 2023. Leveraging Generative AI Solutions in Art and Design Education: Bridging Sustainable Creativity and Fostering Academic Integrity for Innovative Society. *E3S Web of Conferences*. <https://doi.org/10.1051/e3sconf/202342601102>.
- [15] Rania Lampou. 2023. The Integration of Artificial Intelligence in Education: Opportunities and Challenges. *Review of Artificial Intelligence in Education*. <https://doi.org/10.37497/rev.artif.intell.educ.v4i00.15>.
- [16] Anjan Chatterjee. 2022. Art in an age of artificial intelligence. *Frontiers in Psychology*, 13. <https://doi.org/10.3389/fpsyg.2022.1024449>.
- [17] Robert Donmoyer. 1993. *Art Criticism as a Guide to Student Evaluation*. Theory into Practice. <https://www.jstor.org/stable/1476374>
- [18] Theodore F. Wolff and George Geahigan. 1997. *Art criticism and education*. <https://doi.org/10.2307/3333581>.
- [19] Sebastien Fitch. 2016. *Art Critiques: A Guide*. *Studies in Art Education*, 57, 185–187. <https://doi.org/10.1080/00393541.2016.1133207>.
- [20] Jeffrey Broome, Adriane Pereira, and Tom Anderson. 2018. *Critical Thinking: Art Criticism as a Tool for Analysing and Evaluating Art*, *Instructional Practice and Social Justice Issues*, 37, (August 2018), 265–276. <https://doi.org/https://doi.org/10.1111/jade.12111>
- [21] Heidi L. Andrade, Joanna Hefferen, and Maria E. Palma. 2019. *Formative Assessment in the Arts*. In *Routledge eBooks*. 126–145. <https://doi.org/10.4324/9781315166933-6>
- [22] Matthew Bowman. 2020. *The Work of Art Criticism: Collaboration, Communication, Community*. *Arts* 9, 4 (September 2020), 101. <https://doi.org/10.3390/arts9040101>
- [23] Kendall Walton. 2007. *Aesthetics-What? Why? and Wherefore?* *Journal of Aesthetics and Art Criticism* 65, 2 (March 2007), 147–161. <https://doi.org/10.1111/j.1540-594x.2007.00246.x>

- [24] Wenzhong Xu, Jun Meng, S. Kanaga Suba Raja, M. Padma Priya and M. Kiruthiga Devi. 2021. Artificial intelligence in constructing personalized and accurate critique systems for students. *Int. J. Model. Simul. Sci. Comput.*, 14, 2341001:1-2341001:21. <https://doi.org/10.1142/s1793962323410015>.
- [25] Laura Tubino and Chie Adachi. 2022. Developing critique literacy capabilities through an AI automated critique tool. ASCILITE Publications. <https://doi.org/10.14742/apubs.2022.39>.
- [26] Michelene T.H. Chi and Ruth Wylie. 2014. The ICAP Framework: Linking Cognitive Engagement to Active Learning Outcomes. *Educational Psychologist*. 49, 4 (2014), 219–243. <https://doi.org/10.1080/00461520.2014.965823>.
- [27] Ifeanyi G. Ndukwe, Ben K. Daniel and Chukwudi E. Amadi. 2019. A Machine Learning Grading System Using Chatbots. In *Artificial Intelligence in Education. AIED 2019. Lecture Notes in Computer Science*, Vol. 11626. Springer, Cham. https://doi.org/10.1007/978-3-030-23207-8_67.
- [28] Mustafa Fidan and Nurgun Gencel. 2022. Supporting the Instructional Videos With Chatbot and Peer Critique Mechanisms in Online Learning: The Effects on Learning Performance and Intrinsic Motivation. *Journal of Educational Computing Research*, 60(7), 1716–1741. <https://doi.org/10.1177/07356331221077901>.
- [29] Nicholas Leonard. 2020. Entanglement art education: Factoring ARTificial intelligence and nonhumans into future art curricula. *Art Education*, 73(4), 22-28. DOI:10.1080/00043125.2020.1746163
- [30] Yao-Hua Wu. 2015. The Aesthetic and Cultural Significance of Chrysanthemums in Chinese Art. *Arts Journal of Eastern Asia*, 24, 1 (2015), 89-103.
- [31] Nicole M. Radziwill and Morgan C. Benton. 2017. Evaluating Quality of Chatbots and Intelligent Conversational Agents. arXiv: 1704.04579, Retrieved from <http://arxiv.org/abs/1704.04579>
- [32] Mu Fan Ke. 2023. Applications and Challenges of Artificial Intelligence in the Future of Art Education. *Pacific International Journal*, 6(3), 61-65. <https://doi.org/10.55014/pij.v6i3.405>
- [33] My Art Assistant: Fine Art Painting. Retrieved April 1, 2024, from <https://chatgpt.com/g/g-90DXSBvHP-my-art-assistant-fine-art-painting>.
- [34] Art Critic: AI-driven Art Analysis. Retrieved March 29, 2024, from <https://chat.openai.com/g/g-Vc2ia1Zvb-art-critic-ai-driven-art-analysis>.
- [35] Malcolm Budd. 1996. THE AESTHETIC APPRECIATION OF NATURE. *The British Journal of Aesthetics* 36, 3 (July 1996), 207–222. <https://doi.org/10.1093/bjaesthetics/36.3.207>
- [36] Beibei Gao and Ting Cao. 2024. The influence of Chinese painting on contemporary art education in the context of traditional culture. 5, 2 (August 2024), 141–145. <https://doi.org/10.23977/artpl.2024.050221>
- [37] Boom in demand for children’s art education in China – Finn-Sino Association. Retrieved from <https://www.finn-sino.com/boom-in-demand-for-childrens-art-education-in-china/>
- [38] Chang Yu, Xiang Wang, Ada Marie Gallego Mascarinas, Chatchai Rakthin, Nopavan Namtubtim, Yi Shen and Khoirul Anwar. 2024. The Implementation of Art Appreciation Courses in Chinese University General Education: A Case Study. 6, 1 (August 2024), 60–82. <https://doi.org/10.46303/jcsr.2024.5>