



ARTICLE



<https://doi.org/10.1057/s41599-025-05860-2>

OPEN

Effects of AI-generated images in visual art education on students' classroom engagement, self-efficacy and cognitive load

Cunling Bian^{1✉}, Xiaofang Wang², Yingxue Huang², Shan Zhou² & Weigang Lu^{1✉}

In recent years, Generative Artificial Intelligence (GAI) has demonstrated remarkable potential for producing creative content, including artwork, which has sparked discussions regarding its role in education. Although AI-generated art is gaining recognition in the art world, its application in visual art education remains underexplored. To address this gap, we focused on AI-generated images—a specific category within AI-generated art that emphasizes visual representation. We integrated these images into visual art education and conducted an experiment to evaluate their effects. Seventy-eight fifth-grade students were randomly assigned to a treatment group ($n = 39$) and a control group ($n = 39$). Initially, both groups received conventional visual art instruction using classical images. Subsequently, the treatment group was introduced to a GAI-assisted teaching method utilizing AI-generated images, while the control group continued with conventional instruction. The results revealed that students in the treatment group exhibited significantly higher levels of classroom engagement compared to their peers in the control group. Moreover, the treatment group reported a strong sense of self-efficacy with the GAI-assisted method. Importantly, there were no significant differences in cognitive load between the two groups. A comparative analysis of the students' paintings was conducted, focusing on technical skill, adherence to theme, composition and design, creativity and originality, effort, and improvement, which aligns with the increased classroom engagement and self-efficacy, thereby supporting the effectiveness of AI-generated images. This study is one of the pioneering works to propose and validate the use of AI-generated images to address challenges in developing learning materials within the context of visual art education.

¹Ocean University of China, Qingdao, China. ²Qingdao Ningxia Road Primary School, Qingdao, China. ✉email: clbian@ouc.edu.cn; luweigang@ouc.edu.cn

Introduction

Visual art encompasses various forms, such as painting, sketching, and crafting, which facilitate the expression of human emotions and skills. Recognized for its importance, visual art is crucial in developing the competencies of a refined workforce. Consequently, the role of visual art in education is undeniable (Javaheri Pour et al. 2021). Visual art education in schools fosters creativity and contributes to students' overall personality development (Tyler and Likova, 2012). Furthermore, participation in the performing arts enhances fundamental thinking and improves psychological and creative aptitudes (Alter, 2009; Burger and Winner, 2000; Dhanapal et al. 2014). Despite its significance, educators encounter considerable challenges in creating high-quality visual art learning materials. Classical art examples are the most commonly used resources, providing students with a variety of art styles, techniques, and historical contexts that facilitate critical analysis and inspiration. These tangible references enable students to study and emulate established artists, thereby developing their artistic skills and understanding (Eisner, 2003). However, while classical art examples hold value, they may not adequately reflect contemporary practices and technological advancements. For instance, Sullivan emphasizes the necessity of integrating current art practices and digital media into art education to remain relevant (Sullivan, 2010). Additionally, educators often require significant time to research and revise content, resulting in infrequent updates (Eisner, 2004). Therefore, in a rapidly evolving landscape, traditional methods of visual art education struggle to meet the diverse needs of learners seeking a comprehensive understanding of artworks.

GAI is a transformative component of the broader field of artificial intelligence, focused on the creative generation of new content. Building on foundational principles of AI-such as learning from data and identifying patterns-GAI advances this process by synthesizing entirely new data. Its applications in design, simulation, and creative fields are shaping new frontiers in technology and innovation. GAI has achieved significant progress in creativity, enabling computers to generate relevant and original text and images in response to simple natural language prompts (Allingham et al. 2023; Kasneci et al. 2023; Lu et al. 2024; Singhal et al. 2023). Some outputs have become indistinguishable from human creations, even receiving recognition in traditional art contests (Jiang et al. 2023). While the status of AI-generated art remains a controversial topic, with ongoing debates about its artistic legitimacy, its diverse applications highlight its versatility and growing influence in everyday life (Zylinska, 2020). From revolutionizing traditional visual arts to enhancing entertainment, design, and cultural preservation, AI-generated art effectively bridges the gap between technology and human creativity (Epstein et al.; Jin et al. 2024; Kumar et al. 2024). Despite the increasing academic and practical interest in AI-generated art, there has been limited exploration of its potential applications in visual arts education.

In our view, these models encapsulate centuries of human artistic creations, which hold significant relevance for art education. This study focuses on AI-generated images, a specific subset of AI-generated art, and explores their incorporation into visual art education. Specifically, we utilize GAI to produce artwork images for instructional purposes and conduct an experiment to assess the impact of these generated images on students. The study is guided by the following research questions(RQs):

- RQ1: Can AI-generated images enhance students' classroom engagement in visual arts education?
- RQ2: Can AI-generated images improve students' self-efficacy in visual arts education?

- RQ3: Is there any difference in cognitive load between the treatment group receiving AI-generated images and the control group receiving conventional images?
- RQ4: Is there any difference in painting skills between the treatment group receiving AI-generated images and the control group receiving conventional images?

To address these research questions, we conducted a quasi-experimental study involving 78 fifth-grade students, divided into control and treatment groups. The treatment group engaged with a structured GAI-assisted visual art curriculum using Stable Diffusion, while the control group followed conventional methods. Data were collected through validated questionnaires measuring classroom engagement, self-efficacy, and cognitive load, alongside expert evaluations of student artwork. Our findings revealed that the treatment group exhibited significantly higher classroom engagement and self-efficacy, with no increased cognitive load, compared to the control group. These results demonstrate that AI-generated images, when strategically integrated into instruction, can enhance student motivation and creative confidence without overburdening learners. This study contributes an empirical framework for deploying text-to-image models in formal visual art education, highlighting their potential to bridge technological innovation with pedagogical practice.

Literature review

AI-generated art. AI-generated art refers to artwork created with the assistance of artificial intelligence technologies. This paper provides an overview of significant contributions that have shaped the field of AI-generated art. DeepDream generates images based on the representations learned by neural networks, offering insights into how these networks perceive the world and perform classification (Mordvintsev et al. 2015). Although it was not its initial purpose, DeepDream inspired individuals to use AI as a tool for artistic image creation. Additionally, an important early work is Neural Style Transfer, which employs a pretrained convolutional neural network (CNN) to render the semantic content of an image in various styles (Gatys et al. 2016).

Regarding generative models, the Generative Adversarial Network (GAN) represents a milestone in deep learning literature (Goodfellow et al. 2020). ArtGAN is credited with generating interest in the application of GANs for artistic image creation (Tan et al. 2017). Although its output images did not resemble works of the great masters, they effectively captured low-level features of existing artworks. Shortly thereafter, Elgammal et al. (2017) advanced this concept in their paper on Creative Adversarial Networks (CAN), aiming to train a GAN to produce images deemed artistic by the discriminator, while not conforming to any established art styles. Isola et al. (2017) introduced the innovative concept of a conditional GAN, trained on corresponding image pairs. This model has inspired many artists and AI enthusiasts to create artistic images. However, a notable limitation is the requirement for corresponding image pairs, which are not available for all applications. To address this, CycleGAN combines two conditional GANs and cycles between them, enabling the transformation of photographs into styles such as Monet's paintings (Zhu et al. 2017).

Text-to-image models are a category of artificial intelligence that generates images from textual descriptions. DALL-E is a groundbreaking transformer model in this domain (Ramesh et al. 2021). Given a text description, it predicts image tokens and decodes them into an image during inference. While DALL-E generates cartoons and artistic styles effectively, it struggles with accuracy in producing photo-realistic images. Meta AI released a transformer-based text-to-image model that grants users greater

control over generated images by utilizing segmentation maps (Gafni et al. 2022). Today, text-to-image models have transitioned to diffusion models, which produce stunning images and have been shown to outperform GANs in image generation (Dhariwal and Nichol, 2021). OpenAI harnessed this insight for text-to-image generation and introduced GLIDE, a pipeline consisting of a diffusion model for image synthesis and a transformer encoder for text input (Nichol et al. 2022). To enhance text input optimization, Ramesh et al. (2022) employed a prior model to transform text embeddings into CLIP image embeddings before inputting them into the diffusion model. This approach not only improves image quality compared to GLIDE but also enables users to extend the backgrounds of existing images and generate variations. A major advancement in the field is the fully open-source release of Stable Diffusion, notable for its computational efficiency compared to previously mentioned text-to-image models (Rombach et al. 2022). These models can generate highly realistic and complex images from simple text prompts, empowering individuals without programming knowledge to utilize these powerful tools, thereby assisting artists in expressing their creativity and potentially shaping the future of art. In education, incorporating AI-generated art into visual art teaching provides students with a deeper understanding of the intersection between technology and creativity. It encourages exploration of new possibilities, experimentation with innovative techniques, and engagement with contemporary issues in the art world.

AI-generated images are a specific subset of AI-generated art, focusing exclusively on visual content creation (Chen et al. 2024; Ha et al. 2024). These images are produced using text-to-image models, where AI generates visual representations based on text prompts. While AI-generated art encompasses a broader creative spectrum, including music, literature, and multimedia, AI-generated images specifically pertain to the production of 2D visuals such as paintings, photographs, and illustrations.

Generative AI in education. Generative Artificial Intelligence has gained recognition as a powerful tool for transforming education. Yu and Guo (2023) provide a comprehensive overview of the development and technical support of GAI in education, highlighting its significance for research and practice in the field. Among GAI technologies, two main categories are prominent: text-generating and image-generating AI. A notable example of text-generating GAI is ChatGPT (Noy and Zhang, 2023), developed by OpenAI, which utilizes both generative and conversational AI to produce new textual outputs based on user prompts in phrases or sentences (Achiam et al. 2023). Stable Diffusion and DALL-E 2 exemplify image-generating AI, processing user input prompts in text and automatically generating corresponding images.

Several studies have explored the potential and implications of incorporating text-generating GAI into educational settings. Ahmad et al. (2023) emphasize the importance of ChatGPT in the education sector, highlighting its role in learning and teaching reform. Hsu and Ching (2023) further examine the dynamic role of ChatGPT in education, showcasing the evolving landscape of this technology in educational contexts. Conversely, Lawan et al. (2023) propose a modified flipped learning approach to mitigate the adverse effects of ChatGPT on education, addressing implementation concerns. Ruiz-Rojas et al. (2023) focus on the practical application of ChatGPT in education, particularly in instructional design, highlighting the impact of these tools as evidenced by surveys conducted in a MOOC course involving ChatGPT. Wang and Liu (2023) discuss the potential and limitations of ChatGPT and GAI in medical safety education, emphasizing the need to integrate these technologies with real-life

scenarios for comprehensive and personalized medical services. Baytak (2023) conducts a literature review on the acceptance and diffusion of GAI in education, specifically focusing on ChatGPT and Google Bard-type Large Language Models. The review offers insights into current trends and challenges associated with adopting GAI in educational settings. Mao et al. (2024) explore the implications of GAI for assessment in education, highlighting its transformative potential in reshaping traditional assessment methods. Overall, the literature on text-generating GAI in education underscores its transformative potential and the necessity for strategic implementation to effectively harness its benefits.

While ChatGPT has garnered significant attention from learners, researchers, and educators, Stable Diffusion and DALL-E 2 have received less focus in educational contexts. Few studies have investigated the potential benefits of using image-generating AI in education. Dehouche and Dehouche (2023) evaluate the capacity of Stable Diffusion in visual art education, suggesting that these new art creation tools can facilitate the teaching of art history, aesthetics, and technique. Lee et al. (2024) underscore the potential of integrating image-generating AI models in STEAM education, describing a learning activity where students generated creative images with image-generating AI and wrote imaginative diaries inspired by these visuals.

Based on the existing literature, the present study designs a novel GAI-assisted teaching method in which Stable Diffusion serves as the primary tool for generating artwork images used in teaching to enhance visual art education.

Theoretical framework. This study is grounded in constructivist learning theory (Vygotsky, 1978), which posits that learners actively construct knowledge through experimentation, social interaction, and reflection. Generative AI tools align with these principles by enabling iterative exploration of artistic concepts through rapid prototyping, thereby fostering learner agency and creative problem-solving. For instance, GAI's capacity to generate multiple visual variations allows students to test hypotheses about artistic techniques—a process mirroring Bruner's spiral curriculum model (Bruner, 1974), where cyclical refinement deepens conceptual understanding. Meanwhile, cognitive load theory further informs our analysis of how AI tools mediate learning efficiency (Sweller, 1988). By automating repetitive tasks, GAI reduces extraneous cognitive load, enabling students to allocate mental resources to intrinsic load and germane load. This aligns with Chandler and Sweller's findings that well-designed tools optimize cognitive load distribution (Chandler and Sweller, 1991), a dynamic critical to maintaining engagement without overwhelming learners.

Cognitive load in traditional visual art education often arises from the simultaneous demands of technical execution and creative ideation, as evidenced by studies using dual-task paradigms to measure working memory allocation (Paas and Van Merriënboer, 2020; Sweller, 2011). For instance, novice learners exhibit higher extraneous cognitive load when mastering manual techniques like perspective drawing, diverting attention from conceptual development (Huang, 2019). In contrast, GAI tools such as Stable Diffusion reduce procedural demands by automating technical execution, such as rendering complex styles, thereby reallocating cognitive resources toward creative exploration (Pavlik and Pavlik, 2024). However, this shift introduces new germane load components, such as prompt engineering and iterative refinement, which require structured scaffolding to avoid overwhelming learners (Sweller et al. 2019). Recent studies suggest that well-designed GAI interventions can lower intrinsic cognitive load compared to traditional methods, while

Table 1 Demographic information of participants.						
Item	Treatment group(n = 39)		Control group(n = 39)		t	p
	Mean	SD	Mean	SD		
Age	11.18	0.451	11.08	0.422	1.036	0.303
Gender	1.44	0.502	1.51	0.506	0.673	0.503

maintaining equivalent germane load for higher-order skill development (Hwang et al. 2020). This balance positions GAI as a complementary tool rather than a replacement, aligning with principles of multimedia learning, which emphasize optimizing cognitive load to enhance engagement and retention (Mayer, 2002).

Methods

Participants. This study involved 78 fifth-grade students from a primary school in Shandong Province, China. The participants were drawn from two classes: Class A (39 students) and Class B (39 students), comprising 41 males and 37 females. Their ages ranged from 10 to 12 years, with a mean age of 11.13 years and a standard deviation of 0.437 years. While formal assessments of prior AI exposure or artistic skills were not conducted, all participants were drawn from classrooms following Shandong Province’s standardized art and technology curriculum. This ensures parity in foundational art training and introductory AI literacy modules. Teacher interviews further confirmed no extracurricular AI-art programs were offered at the school prior to the study. This homogeneous sampling was intentional to control for variability in socioeconomic and infrastructural factors. While limiting generalizability, this design aligns with design-based research principles for initial theory-building in controlled, real-world settings prior to cross-context validation (Anderson and Shattuck, 2012). Participants were assigned to either the treatment or control group. Importantly, there were no statistically significant differences between the treatment and control groups regarding age and gender ($p > 0.05$; see Table 1). Both groups participated in an art painting course instructed by the fourth author. The course consisted of various modules focusing on different painting skills, with this study concentrating on one specific module aimed at developing students’ cartoon painting skills. This module comprised two 20-minute sessions. In Session 1, explicit instruction on cartoon painting techniques (e.g., facial features and clothing decorations) was provided. In Session 2, students engaged in learning tasks to practice their cartoon painting skills. Students in the treatment group received the standard art curriculum supplemented with structured exposure to AI-generated images, while the students in the control group followed the standard curriculum without AI integration, using traditional resources for the same themes. And both groups shared identical core instruction delivered by the same teacher to minimize instructional variability. Informed consent for participation was obtained prior to the study, and ethical considerations were upheld throughout the design, implementation, and reporting of the research.

Stable diffusion. This study utilized a text-to-image model called Stable Diffusion to generate artworks for visual art education based on text descriptions (Rombach et al. 2022). The model’s code and weights are publicly accessible and can be executed on most consumer hardware. To create images, Stable Diffusion employs CLIP (Radford et al. 2021) to map a text prompt into a joint text-image embedding space, selecting a rough, noisy image that semantically aligns with the input prompt. This image is

subsequently refined using a denoising method grounded in the latent diffusion model to produce the final output. Beyond the text prompt, the Text-to-Image generation script within Stable Diffusion allows users to adjust various parameters, including sampling type, output image dimensions, and seed value.

The functionality of Stable Diffusion for text-to-image generation revolves around the use of prompts as the primary input condition. A prompt guides Stable Diffusion in generating content; therefore, it is advisable to structure the wording systematically. Based on practical experience, this paper outlines a concise generative artwork prompt structure consisting of two components: content description and style description. The content description specifies the elements present in the artwork. To effectively write this part, consider the following questions: What is the main subject? What are its features and details? Are there additional elements aside from the main subject, and how do they relate? What are the features and details of these additional elements? What is the background or environment of the artwork? The style description delineates the artistic style, perspective, and magic words.

For example, the following steps illustrate the content and style descriptions, where step1-step5 belong to content description and step 6 belongs to the style description:

1. **What is the main subject?** To depict a pug dog, we start with the prompt: *a pug dog*.
2. **What are its features and details?** To specify that the pug dog is anthropomorphic and wearing a robe while eating breakfast, we refine the prompt to: *an anthropomorphic pug dog, in a robe, eating breakfast*.
3. **Are there other elements besides the main subject, and what is their relationship to the main subject?** To include a newspaper, we further refine the prompt: *an anthropomorphic pug dog, in a robe, eating breakfast, reading the newspaper*. Here, we establish the relationship between the main subject and the element, i.e., the pug dog is reading the newspaper. It is important to ensure that the relationship between elements is logical.
4. **What are the features and details of these other elements?** To indicate that the newspaper covers the football World Cup, we refine the prompt further: *an anthropomorphic pug dog, in a robe, eating breakfast, reading the newspaper about the Football World Cup*.
5. **What is the background or environment of the artwork?** To set the scene in a kitchen, we add to the prompt: *an anthropomorphic pug dog, in a robe, eating breakfast, reading the newspaper about the Football World Cup, in the kitchen*.
6. **What is the style of the artwork?** To achieve a high-quality cartoon style, we include relevant descriptors: *an anthropomorphic pug dog, in a robe, eating breakfast, reading the newspaper about the Football World Cup, in the kitchen, cartoon character, masterpiece by Disney, high quality*.

The generated images are presented in Fig. 1. The process of generating AI-generated artwork with Stable Diffusion exemplifies the intersection of technology and creativity in visual art education. By leveraging text prompts to guide the model, educators can create tailored artistic content that resonates with students’ interests and learning objectives. The structured approach to developing prompts, as outlined in the preceding sections, ensures that students can articulate their creative visions clearly, facilitating a deeper engagement with the artistic process. This integration of AI-generated images not only enhances the learning experience but also encourages students to explore new forms of artistic expression. Ultimately, the use of Stable

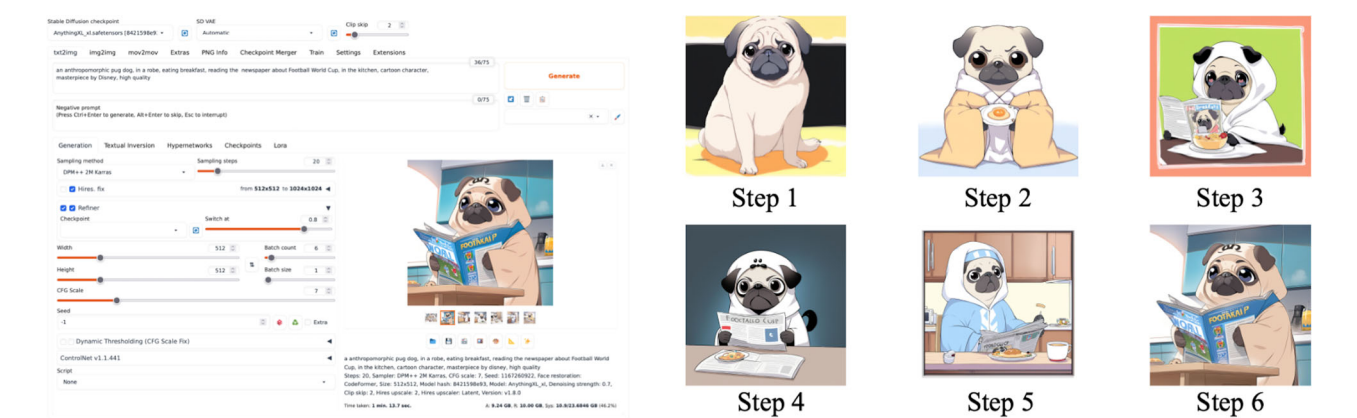


Fig. 1 Text-to-image generation using Stable Diffusion.

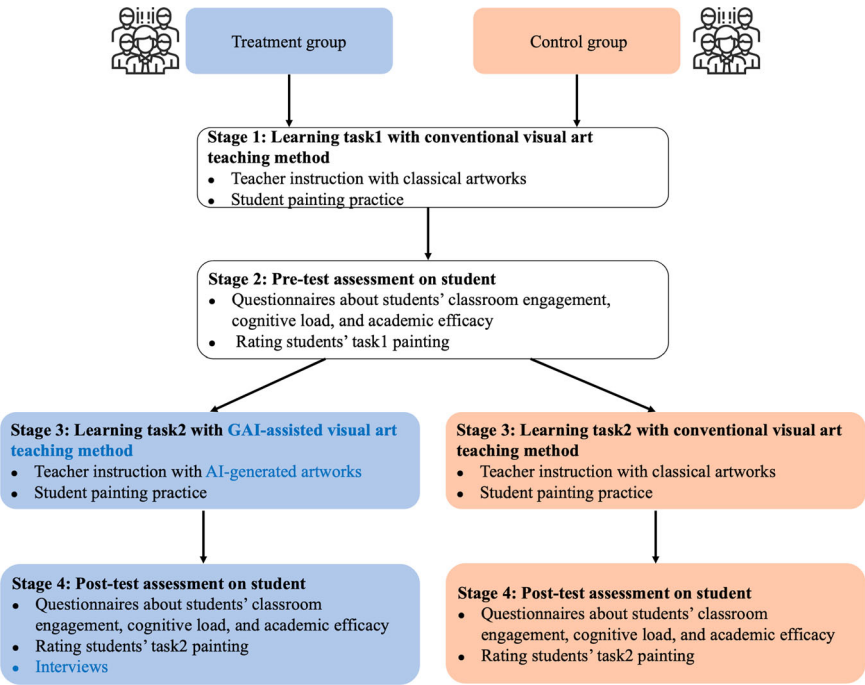


Fig. 2 The process of the experiment.

Diffusion and similar technologies represents a significant advancement in art education, enabling the creation of diverse, innovative artworks that reflect contemporary practices while fostering students' artistic skills and self-efficacy.

Research designs. The experiment comprised four stages, as illustrated in Fig. 2. Stage 1 involved instructing participants on Task 1 using a conventional visual art teaching method. Stage 2 evaluated participants' initial perceptions of visual art education. In Stage 3, the treatment group engaged in a GAI-assisted visual art teaching method for Task 2, while the control group continued with the conventional approach. Stage 4 reassessed participants' perceptions of visual art education. It is important to note that both Task 1 and Task 2 necessitated students' use of similar painting techniques and materials. The instruction and guidance provided were equivalent, and the tasks presented comparable levels of difficulty. The primary distinction between the two visual art teaching methods is the source of the images used in the instructional materials. The GAI-assisted method employs images generated by artificial intelligence (AI-generated images), whereas the conventional method relies on classical artworks. Both

methods follow the same structure and content delivery, ensuring that the only variable affecting students' experiences is the source of the images. This design facilitates a clear comparison of the impact of AI-generated images versus traditional classical art images on the student learning experience.

Stage 1: Learning task 1 with the conventional visual art teaching method

Stage 1 focused on teaching participants the concepts, examples, and applications of Task 1, "Animation Past and Present," utilizing carefully selected classical artworks. These artworks, drawn from iconic cartoon characters, provided targeted resources for students' professional development. This stage consists of two parts: teacher instruction and student painting practice. During the teacher instruction segment, the instructor delivered explicit lessons on the history and principles of animation to both the treatment and control groups, employing well-designed PowerPoint slides featuring selected classical artworks. In the student painting practice segment, students were required to apply the knowledge gained from the lecture in their visual art practice. The task was to "draw animated characters that are vivid, interesting, and rich in color." This assignment

Table 2 The Cronbach's alpha value of questionnaires.		
	Pretest	Posttest
Classroom engagement	0.855	0.866
Self-efficacy	0.858	0.801
Cognitive load	0.822	0.846

closely relates to students’ lives, as many are eager to express their thoughts and skills. The duration of this stage was approximately 40 minutes.

Stage 2: Pre-test assessment on students’ perception
In Stage 2, both the treatment and control groups completed questionnaires assessing their classroom engagement, cognitive load, and academic efficacy within the context of the conventional visual art teaching method. This stage lasted about 8 minutes. To enhance the reliability of responses from primary school students, the questionnaire employed simple, clear language appropriate for their comprehension levels, avoiding complex terminology and lengthy sentences. Additionally, prior to completing the questionnaire, the teacher demonstrated how to fill it out, explaining the meaning of each question and the appropriate response format.

Students’ classroom engagement. Students’ engagement in the conventional visual art teaching method was evaluated using a questionnaire adapted from the research of Wang et al. (2014). This instrument assesses various dimensions of engagement, including attention, interest, investment, and effort expended in visual art learning. It measures five key factors: Affective Engagement, Behavioral Engagement-Compliance, Behavioral Engagement-Effortful Class Participation, Cognitive Engagement, and Disengagement. The questionnaire comprised 24 items and employed a 7-point Likert scale to capture participant responses, ranging from 1 (strongly disagree) to 7 (strongly agree).

Students’ cognitive load. The cognitive load experienced by students during the experiment was evaluated using a questionnaire adapted from the study by Hwang et al. (2013). This instrument focuses on assessing participants’ mental load and the effort required to complete a task. It consists of 8 items and utilizes a 7-point Likert scale for participant responses, ranging from 1 (strongly disagree) to 7 (strongly agree).

Students’ academic efficacy. Students’ academic efficacy in the conventional visual art teaching method was assessed using a questionnaire developed by Midgley et al. (2000), which measures students’ perceptions of their competence in completing class-work. This questionnaire also includes 8 items and employs a 7-point Likert scale to gauge participant responses, from 1 (strongly disagree) to 7 (strongly agree).

The Cronbach’s alpha values for the questionnaires used in this study are presented in Table 2, all demonstrating high internal consistency.

Stage 3: Learning task 2 with GAI-assisted visual art teaching method

Stage 3: Learning task 2 with GAI-assisted visual art teaching method

Moving on to Stage 3, participants engage in visual art learning for Task 2, titled “Anthropomorphic Cartoon Character,” which bears similarities to Stage 1. The key knowledge points for this task include facial features, action characteristics, clothing details, abilities, and expressions. Both groups shared identical core instruction delivered by the same teacher to minimize

instructional variability, following a standardized lesson plan and scripted interactions. The instructor also received explicit training to standardize delivery elements across conditions. This training emphasized adherence to scripted prompts and neutral facilitation strategies to minimize variability unrelated to the intervention. Students in both groups used identical traditional art tools, such as sketching pencils, watercolor paints, and standard art paper, to complete their painting tasks. However, a crucial difference arises in the learning resources employed to create PowerPoint presentations between the treatment and control groups. The treatment group utilizes AI-generated artwork to enhance their learning experience, as illustrated in Fig. 3, while the control group relies on classical artworks selected from iconic cartoon characters. Both sets of examples were curated to align with the learning objectives of Task 2 and matched in thematic complexity, artistic style, and pedagogical relevance. A full list of classical artworks used in the control group is provided in Fig. 4. This ensured that the only variable differentiating the groups was the source of instructional imagery. This stage consists of two components: teacher instruction and student painting practice. The objective of the painting practice is as follows: “Using the anthropomorphism techniques you have learned, create and design a distinct and characteristic anthropomorphic cartoon character.” This stage lasts approximately 40 minutes.

Stage 4: Post-test assessment on students’ perception
In the final stage, both treatment and control groups complete questionnaires assessing their classroom engagement, cognitive load, and academic efficacy in Task 2. Additionally, the paintings created during the task are evaluated for all participants. Three professional visual art teachers are tasked with this evaluation, using a detailed rubric that considers both the technical and expressive dimensions of the students’ work. To mitigate evaluation bias, raters were blinded to group assignments during artwork scoring. Artworks were anonymized and randomized, with no identifiers linking them to treatment/control conditions. Each painting is assessed according to the rubric criteria, with scores ranging from 1 to 5 points based on how well each criterion is met. The rubric includes: (1) Technical Skill: How effectively did the student apply the techniques taught in class? (2) Adherence to Theme: Did the student follow the assignment guidelines or theme? (3) Composition and Design: How effectively is space, balance, and harmony utilized? (4) Creativity and Originality: How unique and imaginative is the artwork? (5) Effort and Improvement: How much effort did the student invest, and have they demonstrated progress over time?

Data analysis. To investigate the effects of AI-generated artworks on students’ classroom engagement, self-efficacy, and cognitive load in visual art education (RQ1, RQ2, and RQ3), independent samples t-tests and analysis of covariance (ANCOVA) were conducted. The t-test aimed to identify differences in classroom engagement, self-efficacy, and cognitive load between the control and treatment groups during the pretest phase, as they are optimal for detecting mean differences in continuous variables. The ANCOVA was subsequently applied to posttest outcomes to statistically control for pretest scores as covariates, thereby isolating the intervention’s effect while adjusting for initial variability. Notably, the data collected during the rating of students’ paintings met the assumptions necessary for conducting t-tests and ANCOVA. To explore the impact of AI-generated images on students’ painting skills (RQ4), independent sample t-tests and ANCOVA were performed to examine differences between the treatment and control groups.






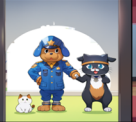



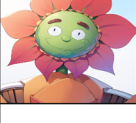






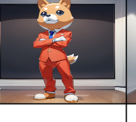

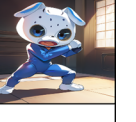
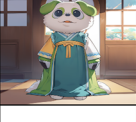



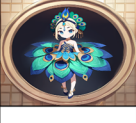





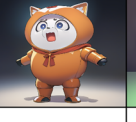


Theme	Model	AI-generated Artworks				
Leading-in	realvisxIV40_v40Lig htningBakedvae					
		A panda doctor, anthropomorphic, masterpiece, best quality, ultra-detailed, photorealistic, realistic, photography	A elephant mayor, anthropomorphic, masterpiece, best quality, ultra-detailed, photorealistic, realistic, photography	A anthropomorphic warrior cheetah, cheetah in loose red, white robes, a long straight wooden staff, katana blades in his belt	A anthropomorphic hedgehog cooking in the kitchen, lightroom, cinematic, hdr, raw, intricate, high quality, soothing tones, intricate details, low contrast	A anthropomorphic pug dog in a robe, eating breakfast, reading the newspaper, lightroom, cinematic, hdr, raw, intricate, high quality, soothing tones, intricate details, low contrast
	AnythingXL_xl					
		Many anthropomorphic animals, Zoo, dog, cat, bird, tree, flower, grass, surprised, face, full body, two eyes, eyebrow, nose, mouth, ear, cartoon character, master piece, high quality	A anthropomorphic dog, a anthropomorphic cat, police suit, shake hands, face, full body, two eyes, eyebrow, nose, mouth, ear, cartoon character, master piece, high quality	(Human:1.6), act like a fox, full body, two eyes, eyebrow, nose, mouth, ear, cartoon character, master piece, high quality	(Human:1.6), act like a fox, full body, two eyes, eyebrow, nose, mouth, ear, cartoon character, master piece, high quality	(Human:1.6), act like a fox, full body, two eyes, eyebrow, nose, mouth, ear, cartoon character, master piece, high quality
Facial Feature	cartoonArcadiaSDX LSD1_xenoArcadiaC X					
		(A anthropomorphic flower:1.6), face, eye, eyebrow, nose, mouth, ear, cartoon character, master piece, high quality	A anthropomorphic hamburger, cute	A anthropomorphic apple, cute, cartoon character	(A anthropomorphic police car:1.6), police suit, wheel, face, full body, two eyes, eyebrow, nose, mouth, ear, cartoon character, master piece, high quality	(A anthropomorphic mobile phone:1.6), suit, wheel, face, two eyes, eyebrow, nose, mouth, ear, cartoon character, master piece, high quality
Action Feature	AnythingXL_xl					
		(A anthropomorphic dog:1.6), running on street, face, full body, eye, eyebrow, nose, mouth, ear, cartoon character, master piece, high quality	(A anthropomorphic dog:1.6), hand wave, face, full body, two eyes, eyebrow, nose, mouth, ear, cartoon character, master piece, high quality	(A anthropomorphic dog:1.6), crossed arms, face, full body, eye, eyebrow, nose, mouth, ear, cartoon character, master piece, high quality	(A anthropomorphic dog:1.6), stand, face, full body, eye, eyebrow, nose, mouth, ear, cartoon character, master piece, high quality	(A anthropomorphic dog:1.6), fighting stance, face, full body, two eyes, eyebrow, nose, mouth, ear, cartoon character, master piece, high quality
Clothing Decoration	AnythingXL_xl					
		(A anthropomorphic dog:1.6), hanfu, face, full body, two eyes, eyebrow, nose, mouth, ear, cartoon character, master piece, high quality	(A anthropomorphic dog:1.6), tianra, face, full body, two eyes, eyebrow, nose, mouth, ear, cartoon character, master piece, high quality	(A anthropomorphic dog:1.6), firefighter jacket, face, full body, two eyes, eyebrow, nose, mouth, ear, cartoon character, master piece, high quality	(A anthropomorphic dog:1.6), China school uniform, face, full body, two eyes, eyebrow, nose, mouth, ear, cartoon character, master piece, high quality	(A anthropomorphic dog:1.6), hardhat, face, full body, two eyes, eyebrow, nose, mouth, ear, cartoon character, master piece, high quality
Ability Feature	AnythingXL_xl					
		(Human:1.6), act like a peacock, full body, two eyes, eyebrow, nose, mouth, ear, cartoon character, master piece, high quality	(Human:1.6), act like a butterfly, full body, two eyes, eyebrow, nose, mouth, ear, cartoon character, master piece, high quality	(Human:1.6), act like a monkey, full body, two eyes, eyebrow, nose, mouth, ear, cartoon character, master piece, high quality	A humanoid body with distinct human features, strong body and long tusks, jacket, full body, two eyes, eyebrow, nose, mouth, ear, cartoon character, master piece, high quality	Boy, a humanoid body with distinct human features, bravery and power of lion, full body, two eyes, eyebrow, nose, mouth, ear, cartoon character, master piece, high quality
Expression Feature	AnythingXL_xl					
		(A anthropomorphic dog:1.6), happy, pleased, face, full body, two eyes, eyebrow, nose, mouth, ear, cartoon character, master piece, high quality	(A anthropomorphic dog:1.6), sad, face, full body, two eyes, eyebrow, nose, mouth, ear, cartoon character, master piece, high quality	(A anthropomorphic dog:1.6), scared, face, full body, two eyes, eyebrow, nose, mouth, ear, cartoon character, master piece, high quality	(A anthropomorphic dog:1.6), disgust, face, full body, two eyes, eyebrow, nose, mouth, ear, cartoon character, master piece, high quality	(A anthropomorphic dog:1.6), surprised, face, full body, two eyes, eyebrow, nose, mouth, ear, cartoon character, master piece, high quality

Fig. 3 Text-to-image generation examples.

Results

Classroom engagement. An independent samples t-test was initially performed to analyze classroom engagement scores from the pretest. As shown in Table 3, the treatment group had a mean score of 28.385 (SD = 3.066), while the control group had a mean

score of 28.667 (SD = 2.388). The t-test results ($t = 0.453$, $p = 0.652$) indicated no significant differences in affective engagement between the two groups during the pretest. Comparable results were observed in the dimensions of behavioral engagement, cognitive engagement, and disengagement.



Fig. 4 Classical artworks selected from iconic cartoon characters.

Table 3 The t-test value of classroom engagement in the pretest.						
	Group	n	Mean	SD	t	p
Affective engagement	Treatment	39	28.385	3.066	0.453	0.652
	Control	39	28.667	2.388		
Behavioral engagement	Treatment	39	40.692	8.733	0.815	0.418
	Control	39	39.436	4.057		
Cognitive engagement	Treatment	39	37.410	4.865	0.423	0.673
	Control	39	37.872	4.770		
Disengagement	Treatment	39	8.000	2.162	1.075	0.286
	Control	39	7.564	1.318		

Following this, an ANCOVA was conducted to investigate differences between the groups, using the pretest scores as a covariate and the posttest scores as the dependent variable. Table 4 presents the adjusted mean trust scores, with the treatment group scoring 31.522 (SD = 2.426) and the control group scoring 26.504 (SD = 2.945) in the posttest. The ANCOVA revealed a significant difference between the treatment and control groups ($F = 17.279, p < 0.001$), indicating that students receiving the GAI-assisted visual art teaching method exhibited significantly higher levels of affective engagement in the classroom compared to those receiving the conventional teaching method. Similar findings were noted in the dimensions of behavioral and cognitive engagement. Regarding disengagement, the adjusted mean trust scores in the posttest were 5.354 (SD = 1.775) for the treatment group and 7.979 (SD = 2.211) for the control group. The ANCOVA also indicated a significant difference between the two groups ($F = 54.482, p < 0.001$), demonstrating that the treatment group showed significantly lower levels of disengagement compared to the control group. The pretest analysis confirmed baseline equivalence between groups across all classroom engagement dimensions. Post-intervention ANCOVA results revealed substantial improvements in the treatment group compared to the control group.

Self-efficacy. An independent samples t-test was initially performed on self-efficacy scores obtained during the pretest. As shown in Table 5, the treatment group exhibited a mean self-efficacy score of 24.846 (SD = 2.033), while the control group had a mean score of 24.667 (SD = 2.527). The t-test results ($t = 0.346, p = 0.731$) indicated no significant differences in self-efficacy between the two groups in the visual art courses during the pretest.

Following this, an ANCOVA was conducted to assess the differences between the groups, using the pretest scores as a covariate and the posttest scores as the dependent variable. As presented in Table 6, the adjusted mean self-efficacy scores in the posttest were 27.615 (SD = 2.347) for the treatment group and 24.846 (SD = 1.387) for the control group. The ANCOVA results revealed a significant difference between the two groups ($F = 34.242, p < 0.001$), indicating that students who received the GAI-assisted visual art teaching method demonstrated significantly higher levels of self-efficacy compared to those who received the conventional teaching method. While pretest self-efficacy scores were equivalent, the treatment group showed an increase in posttest scores. This suggests that GAI-assisted instruction not only enhanced technical skills but also bolstered students' confidence in their artistic abilities.

Cognitive load. To evaluate the difference in cognitive load between the treatment and control groups, an independent samples t-test was performed. As presented in Table 7, the treatment group achieved a mean score of 17.513 (SD = 6.890), while the control group had a mean score of 16.667 (SD = 6.225). The t-test results ($t = 0.569, p = 0.571$) revealed no significant differences between the two groups, indicating that the cognitive load experienced by the treatment group was similar to that of the control group throughout the experiment. Despite the added complexity of GAI integration, cognitive load remained comparable between groups, indicating that the intervention did not overwhelm students. This supports the feasibility of implementing GAI tools in classroom settings without undue mental strain.

Visual art work. A comprehensive quality analysis of the students' visual artworks was conducted. Examples of these artworks are presented in Fig. 5. Three primary school teachers served as raters for the evaluation. As shown in Table 8, Rater Z, who has 24 years of experience teaching art, is the most senior participant and also the instructor for this experiment. Rater H has taught art for four years, while Rater M has three years of experience teaching mathematics. The inclusion of a math teacher in the rating process was intentional; she has a longstanding commitment to the practice of art, providing a unique perspective on observation and analysis. Additionally, her expertise enables her to offer professional insights into aspects such as structure, symmetry, proportion, and geometric aesthetics of the artworks. Rater Z holds a college degree, while the other participants possess bachelor's degrees. Consent was obtained from all raters to participate in this study.

An independent samples t-test was performed on the scores of the students' artworks. As outlined in Table 9, the treatment group achieved a mean score of 4.487 (SD = 0.389), compared to

Table 4 The ANCOVA value of classroom engagement in the posttest.							
	Group	n	Mean	SD	Adjusted mean	F	p
Affective engagement	Treatment	39	31.462	2.426	31.522	17.379	<0.001
	Control	39	26.564	2.945	26.504		
Behavioral engagement	Treatment	39	42.641	6.239	42.245	190.901	<0.001
	Control	39	39.282	3.539	39.697		
Cognitive engagement	Treatment	39	42.128	3.629	42.212	20.824	<0.001
	Control	39	36.769	3.842	36.686		
Disengagement	Treatment	39	5.523	1.775	5.354	54.482	<0.001
	Control	39	7.821	2.211	7.979		

Table 5 The t-test value of Self-efficacy in the pretest.					
Group	n	Mean	SD	t	p
Treatment	39	24.846	2.033	0.346	0.731
Control	39	24.667	2.527		

Table 7 The t-test value of Cognitive load.					
Group	n	Mean	SD	t	p
Treatment	39	17.513	6.890	0.569	0.571
Control	39	16.667	6.225		

Table 6 The ANCOVA value of Self-efficacy in the posttest.						
Group	n	Mean	SD	Adjusted mean	F	p
Treatment	39	27.615	2.347	27.658	34.242	<0.001
Control	39	24.846	1.387	24.804		

the control group’s mean score of 4.248 (SD = 0.437). The t-test analysis ($t = 2.554$, $p = 0.006$) revealed significant differences between the two groups, indicating that the treatment group exhibited higher levels of technical skill than the control group throughout the experiment. Similar results were observed in the dimensions of adherence to theme, composition and design, creativity and originality, and effort and improvement. To ensure the reliability of the artwork evaluations, inter-rater consistency was assessed using the Intraclass Correlation Coefficient (ICC) with a two-way mixed-effects model for absolute agreement. The ICC value for the composite score,averaged across all rubric criteria, was 0.847 ($p < 0.001$), indicating excellent agreement among the three raters.

Discussion

Artificial Intelligence has made significant advancements in the realm of creativity. Researchers have utilized Generative AI to produce original visual art images in response to simple natural language prompts. However, the potential application of AI-generated art within visual arts education has received limited attention, despite its capacity to foster creativity and allow students to experiment with diverse artistic concepts beyond the constraints of traditional materials. To explore this potential, we integrated AI-generated images into visual art education and assessed their impact on students. We developed a structured prompt format for generative artwork that included content and style descriptions. Using Stable Diffusion, we generated images for visual art instruction. To evaluate the effectiveness of these AI-generated images, we recruited 78 fifth-grade students, assigning them to either a control group or a treatment group. Initially, both groups learned through conventional visual art teaching methods, but the treatment group subsequently engaged with GAI-assisted visual art education. Data were collected through questionnaires and students’ paintings, yielding valuable insights.

Firstly, the treatment group demonstrated a significantly higher level of affective engagement with the AI-generated images used

in visual art education, encompassing both behavioral and cognitive engagement dimensions. This heightened emotional involvement can be attributed to the novelty and creativity of the AI-generated images, which capture students’ attention and stimulate their interest, ultimately enhancing the overall learning experience. Prior research suggests that innovative educational materials can elevate emotional engagement by triggering curiosity and interest (Chen and Wu, 2015; Mayer, 2002; Renninger and Hidi, 2015). The interactive nature of working with GAI promotes active participation and hands-on learning, thereby improving students’ behavioral engagement. By involving students in the image generation process through AI, they become more engaged in the lesson and are likely to remain focused on the task. This finding fosters a dynamic classroom environment, where students are more inclined to participate actively. Moreover, GAI enhances cognitive engagement by stimulating critical thinking and creative exploration. By allowing teachers to create and manipulate AI-generated images, they can nurture creativity and encourage students to think critically about the subject matter. For instance, students can investigate various historical art styles by generating images in those styles, leading to a deeper understanding of artistic movements and their cultural contexts. A focus on students’ needs, interests, and active participation renders learning more relevant and personalized. When students perceive that the material is tailored to their experiences and goals, they are more likely to engage with the content, participate in discussions, and take ownership of their learning (Diwan et al. 2023; Lambert et al. 2017; Walsh et al. 2021). These outcomes were anticipated, as learner-centered content represents a significant advantage of AI-generated methods over traditional approaches. The heightened affective engagement aligns with experiential learning principles, where tools bridging abstract concepts and practical application enhance intrinsic motivation (Dewey, 1986). The novelty of AI-generated images mirrors historical shifts in education, such as the introduction of multimedia tools, which similarly leveraged interactivity to trigger curiosity. Furthermore, the behavioral engagement gains resonate with Ryan and Deci’s self-determination theory, where autonomy and competence drive active participation (Deci and Ryan, 2012). This parallels the adoption of early digital tools in fostering learner-centered environments (Papert, 2020).

Secondly, in comparison to the control group that received conventional visual art instruction, the treatment group utilizing GAI-assisted methods exhibited a significantly higher level of self-efficacy. When students encounter content that resonates with



Fig. 5 Examples of students' visual art work.

Table 8 Profile of the participating raters.					
Rater	Gender	Age	Years of teaching	Educational background	Teaching subject
Z	Female	42	24	College degree	Art
H	Female	26	4	Bachelor's degree	Art
M	Female	28	3	Bachelor's degree	Math

their interests or is appropriately adapted to their learning levels, they are more likely to feel confident in their ability to comprehend and apply the material. Additionally, AI-generated images tend to be more dynamic, diverse, and visually engaging than traditional images. This variety captures students' attention and helps sustain their interest, fostering a sense of competence in mastering the subject matter (Keller, 1987). AI-generated images can be crafted to minimize unnecessary details and emphasize key concepts, thereby reducing cognitive load. This simplification enables students to process information more efficiently, enhancing their confidence in their learning capabilities (Uswatun et al. 2020). For instance, Margolis and McCabe (2003) demonstrated that reducing complexity in the learning process aids struggling learners in building confidence. By highlighting essential elements and eliminating distractions, AI-generated images enable students to concentrate on what is most important, enhancing their competence and assurance throughout their learning journey. Similarly, Kim et al. (2025) found that the perceived availability of

AI support significantly boosts students' self-efficacy in task achievement. The self-efficacy improvements reflect Bandura's social cognitive theory, where mastery experiences strengthen learners' agency (Bandura et al. 1999). The reduction of cognitive load through simplified AI-generated visuals echoes Sweller's Cognitive Load Theory, enabling students to allocate mental resources to higher-order tasks (Sweller, 1988). This mirrors the impact of graphing calculators in mathematics education, which reduced computational burdens and redirected focus to problem-solving (Ellington, 2006).

Thirdly, no significant difference was observed in cognitive load between the control and treatment groups. The lack of significant differences in cognitive load reflects the intentional parity in task complexity between the control and treatment groups. Both groups engaged in activities requiring equivalent intrinsic cognitive demands, ensuring a fair comparison of instructional methods. Despite the differences in teaching methods, the fundamental cognitive demands on students remained consistent. In this study, GAI was introduced in a user-friendly manner, ensuring that participants did not experience increased cognitive load compared to the control group. The GAI-assisted method was designed to complement and enhance existing teaching strategies without adding to cognitive demands. By clarifying concepts without introducing new complexities (Avello et al. 2024), it maintained a cognitive load similar to that of the conventional approach. This notion aligns with findings from Huang et al. (2024), which indicated that AI-assisted instructional strategies help reduce cognitive load by systematically guiding students, enabling them to engage with learning content without

Table 9 The t-test value of student work.						
	Group	n	Mean	SD	t	p
Technical Skill	Treatment	39	4.487	0.389	2.554	0.006
	Control	39	4.248	0.437		
Adherence to Theme	Treatment	39	4.615	0.379	2.550	0.006
	Control	39	4.376	0.447		
Composition and Design	Treatment	39	4.530	0.437	2.725	0.004
	Control	39	4.248	0.476		
Creativity and Originality	Treatment	39	4.581	0.444	3.515	<0.001
	Control	39	4.205	0.499		
Effort and Improvement	Treatment	39	4.624	0.352	3.502	<0.001
	Control	39	4.325	0.401		

aimless contemplation. The absence of cognitive load differences supports Vygotsky’s concept of technology as a mediator within the Zone of Proximal Development, where tools scaffold learning without overwhelming learners (Vygotsky, 1978). Similar outcomes were observed with word processors in writing instruction, which minimized mechanical burdens to prioritize creative expression (Haas, 2013). GAI tools replicate this dynamic, balancing innovation with pedagogical intentionality.

Fourthly, students in the treatment group expressed high levels of satisfaction with their visual artworks, evaluating aspects such as technical skill, adherence to theme, composition and design, creativity and originality, effort and improvement. This aligns with the increased classroom engagement and self-efficacy observed, further supporting the effectiveness of the GAI-assisted visual art teaching method. GAI-assisted methods empower students to experiment with various techniques and styles without the limitations imposed by traditional materials. The ability to iterate rapidly and explore multiple approaches contributes to a deeper understanding of technical skills. This phenomenon has also been observed in other studies (Luong and Tran, 2024; Mulian et al. 2024). GAI helps students maintain thematic consistency by providing references and inspiration aligned with their chosen theme, leading to a more cohesive and satisfying artistic experience. Additionally, GAI can guide students in understanding composition and design principles by offering visual templates or suggestions, facilitating the creation of more balanced and aesthetically pleasing compositions, which boosts satisfaction with the final results. GAI-assisted methods stimulate creativity by providing diverse and innovative ideas that students might not have otherwise considered. This inspiration leads to more original and innovative artworks, enhancing students’ satisfaction with their creative accomplishments. Previous literature on GAI indicates that its effects on creativity are complex and multifaceted. On one hand, GAI can significantly enhance individuals’ innovative capabilities, particularly for those with lower creativity (Doshi and Hauser, 2024; Noy and Zhang, 2023; Zhou and Lee, 2024). However, such technological involvement may also introduce risks, including the homogenization of creative content and over-reliance on technology (Jia et al. 2024; Marrone et al. 2024). The efficiency of GAI tools allows students to focus their efforts on areas requiring improvement, accelerating their learning progress. The visible advancements resulting from concentrated effort lead to greater satisfaction. Overall, GAI-assisted visual art teaching methods contribute to a more fulfilling and rewarding learning experience. The high satisfaction with AI-assisted artworks aligns with Bruner’s spiral curriculum model, which emphasizes iterative exploration and

refinement through cyclical learning (Bruner, 1974). GAI enables rapid prototyping allowing students to experiment freely with diverse techniques before refining outputs through traditional methods. This mirrors historical shifts in art education, such as the integration of digital tools like Photoshop, which complemented manual techniques by enabling new forms of creative iteration (Peppler, 2010). The balance between AI-enabled exploration and teacher-guided refinement reflects Mishra and Koehler’s Technological Pedagogical Content Knowledge framework, where effective technology integration hinges on aligning tools with pedagogical goals and subject-specific expertise (Mishra and Koehler, 2006). For instance, educators might use GAI-generated templates to spark student experimentation, then employ traditional critiques to deepen understanding of compositional principles—a synergy that leverages AI’s efficiency while preserving the irreplaceable role of human mentorship in fostering artistic mastery.

Implications, limitations and future directions

Our findings indicate the promising potential of AI-generated image tools, such as Stable Diffusion, in enhancing visual art education. We recommend that educators integrate the generation capabilities of Stable Diffusion with their instructional designs. In this process, teachers can leverage the strengths of Stable Diffusion while also addressing its limitations with their own expertise. For instance, Stable Diffusion can rapidly produce a large volume of artworks, allowing educators to be more discerning in selecting learning materials. They can assess, choose, and adopt AI-generated images based on their suitability for their students. During this selection process, teachers should draw on their understanding of the class dynamics and individual student needs. Although it may seem efficient to provide students with AI-generated images directly, we suggest that educators first critically evaluate the outputs from Stable Diffusion, make necessary adjustments, and incorporate their own concepts before presenting the final learning materials to their students.

As one of the first attempts to examine AI-generated images’ potential in visual arts education settings, the present study has some noteworthy limitations that may open up avenues for future research. First, we used anthropomorphic cartoon characters as our teaching targets. Arguably, AI-generated images on different painting genres may vary (Cao et al. 2023). As a result, it is necessary to examine Stable Diffusion’s outputs on other genres, such as Chinese brush painting and oil painting, and compare its feedback with that provided by teachers to obtain a complete picture of AI-generated images’s ability and to seek the optimal usage pattern. Moreover, we only compared the quantity and type of feedback provided by AI-generated images and by conventional images. Future studies may further examine and compare their quality, such as effectiveness (do they correctly follow aesthetic rules or highlight teaching knowledge points?), by involving experts (Hitsuwari et al. 2023).

Secondly, while our study primarily examined students’ perceptions of AI-generated images in visual arts education, we advocate for future research to include teachers’ perspectives as well. Additionally, the quality of AI-generated images is significant, but their effectiveness ultimately hinges on students’ ability to interpret and utilize these images to enhance their painting skills (Kwegyiriba et al. 2022). It would be beneficial to explore how students’ levels of aesthetic appreciation and painting proficiency influence their perceptions and use of AI-generated images. Future studies should consider these factors to gain deeper insights into optimizing AI-generated images for visual arts education.

Thirdly, our study was limited to 78 fifth-grade students from a primary school in Shandong Province. While this sample provided valuable insights into younger learners' interactions with AI-generated images, students' perceptions and use of technology may vary across grade levels, educational systems, and cultural contexts. For instance, Staddon observed that older students often use technology less frequently than younger students, despite greater exposure over time (Staddon, 2020). Similarly, regional disparities and curricular frameworks could influence how AI tools are adopted and perceived (Lyu et al. 2019; Wu et al. 2020). Future studies should include multi-grade cohorts and diverse educational settings to assess whether the observed benefits of AI-generated images generalize across developmental stages and cultural contexts. For example, comparative studies could examine how pedagogical strategies for AI integration differ between primary and secondary education or between Eastern and Western art curricula.

Fourth, because of the access issue, the participating teachers and students could not independently utilize ChatGPT to generate images for visual arts education; instead, a research assistant obtained the AI-generated images for them. In-depth exploration of teacher-AI collaboration is critical, as teachers' pedagogical expertise significantly impacts the effective use of AI tools in educational settings (Jeon and Lee, 2023). Future research should involve teachers directly using Stable Diffusion and investigating their strategies for interacting with the tool (e.g., the prompts employed to generate outputs) for visual arts education. Such studies will provide valuable insights into the dynamics of collaboration between educators and Stable Diffusion. Additionally, existing literature highlights the benefits of generative AI, including time savings, access to a wide range of information, personalized tutoring, and improved learning retention (Ngo, 2023; Skjuve et al. 2023; Zhang, 2023). Encouraging students to directly engage with Stable Diffusion could allow for image generation tailored to specific educational needs or preferences, making learning more relevant and personalized. For instance, students could create visual representations of their narratives or ideas, aligning with their interests and learning styles.

Fifthly, we utilized specific prompts to elicit Stable Diffusion's output images concerning content and style. Notably, Stable Diffusion is sensitive to variations in input phrasing (Borji, 2022; Du et al. 2024; Mahajan et al. 2024). Changes in prompts can lead to significantly different outputs. Future research should emphasize the importance of prompt programming, experimenting with diverse inputs, and comparing their output results to identify effective strategies for utilizing Stable Diffusion in visual art generation. For example, providing Stable Diffusion with background information about students, such as their personalities and painting skills, may enhance the personalization of the generated images. Researchers should explore prompts that incorporate such information to assess their effectiveness.

Finally, it is important to note that since the introduction of Stable Diffusion, numerous diffusion-based image generation tools have emerged. However, our study focused specifically on Stable Diffusion due to its widespread recognition and significance in the field. Future research could conduct comparative analyses of various diffusion-based tools to identify their respective strengths and weaknesses in facilitating image generation for visual arts education. Such comparisons would be beneficial for educators and researchers alike. Furthermore, since the drafting of this paper, advanced diffusion-based tools, such as Midjourney, Fooocus, DALL-E 3, and FLUX, have become available (Reddy et al. 2021; Ruskov, 2023). Thus, there is an opportunity for subsequent studies to build upon our findings and conduct more comprehensive investigations in this area from a broader perspective.

Data availability

No datasets were generated or analysed during the current study.

Received: 12 November 2024; Accepted: 30 August 2025;

Published online: 30 September 2025

References

- Achiam J, Adler S, Agarwal S, Ahmad L, Akkaya I, Aleman FL, Almeida D, Altenschmidt J, Altman S, Anadkat S (2023) Gpt-4 technical report. arXiv preprint arXiv:2303.08774
- Avello R, Gajderowicz T, Gómez-Rodríguez VG (2024) Is ChatGPT helpful for graduate students in acquiring knowledge about digital storytelling and reducing their cognitive load? an experiment. *Rev Educ Dist* 24 (78)
- Alter F (2009) Understanding the role of critical and creative thinking in Australian primary school visual arts education. *Int Art Early Child Res J* 1(1):1–12
- Ahmad N, Murugesan S, Kshetri N (2023) Generative artificial intelligence and the education sector. *Computer* 56(6):72–76
- Allingham JU, Ren J, Dusenberry MW, Gu X, Cui Y, Tran D, Liu JZ, Lakshminarayanan B (2023) A simple zero-shot prompt weighting technique to improve prompt ensembling in text-image models. In: *Proceedings of the International Conference on Machine Learning*
- Anderson T, Shattuck J (2012) Design-based research: A decade of progress in education research? *Educ Res* 41(1):16–25
- Baytak A (2023) The acceptance and diffusion of generative artificial intelligence in education: A literature review. *Curr Perspect Educ Res* 6(1):7–18
- Bandura A, Freeman WH, Lightsey R (1999) Self-efficacy: The exercise of control. *J Cogn Psychother* 13(2):158–166
- Borji A (2022) Generated faces in the wild: Quantitative comparison of stable diffusion, midjourney and DALL-E 2. arXiv preprint arXiv:2210.00586
- Bruner JS (1974) *Toward a Theory of Instruction*. Harvard University Press, London
- Burger K, Winner E (2000) Instruction in visual art: Can it help children learn to read? *J Aesthet Educ* 34(3/4):277–293
- Chen J, An J, Lyu H, Kanan C, Luo J (2024) Learning to evaluate the artness of AI-generated images. *IEEE Trans Multimed* 26:10731–10740
- Cao Y, Li S, Liu Y, Yan Z, Dai Y, Yu PS, Sun L (2023) A comprehensive survey of AI-generated content (AIGC): A history of generative AI from GAN to ChatGPT. arXiv preprint arXiv:2303.04226
- Chandler P, Sweller J (1991) Cognitive load theory and the format of instruction. *Cogn Instr* 8(4):293–332
- Chen C-M, Wu C-H (2015) Effects of different video lecture types on sustained attention, emotion, cognitive load, and learning performance. *Comput Educ* 80:108–121
- Dehouche N, Dehouche K (2023) What's in a text-to-image prompt? The potential of stable diffusion in visual arts education. *Heliyon* 9(6):16757
- Dewey J (1986) Experience and education. *Educ Forum* 50(3):241–252
- Doshi AR, Hauser OP (2024) Generative AI enhances individual creativity but reduces the collective diversity of novel content. *Sci Adv* 10(28):5290
- Dhanapal S, Kanapathy R, Mastan J (2014) A study to understand the role of visual arts in the teaching and learning of science. In: *Asia-Pacific Forum on Science Learning & Teaching*, vol. 15, pp. 1–25
- Du C, Li Y, Qiu Z, Xu C (2024) Stable diffusion is unstable. In: *Proceedings of the Advances in Neural Information Processing Systems*
- Dhariwal P, Nichol A (2021) Diffusion models beat GANs on image synthesis. In: *Proceedings of the Advances in Neural Information Processing Systems*
- Deci EL, Ryan RM (2012) Self-determination theory. *Handb Theories Soc Psychol* 1(20):416–436
- Diwan C, Srinivasa S, Suri G, Agarwal S, Ram P (2023) AI-based learning content generation and learning pathway augmentation to increase learner engagement. *Comput Educ: Artif Intell* 4:100110
- Epstein Z, Hertzmann A, Herman L, Mahari R, Frank M, Groh M, Schroeder H, Smith A, Akten M, Fjeld J (2023) Art and the science of generative AI: A deeper dive. arXiv 2023. arXiv preprint arXiv:2306.04141
- Eisner EW (2003) The arts and the creation of mind. *Lang Arts* 80(5):340–344
- Eisner EW (2004) What can education learn from the arts about the practice of education? *Int J Educ Arts* 5(4):1–13
- Elgammal A., Liu B, Elhoseiny M, Mazzone M (2017) Can: Creative adversarial networks generating “art” by learning about styles and deviating from style norms. In: *Proceedings of the International Conference on Computational Creativity*
- Ellington AJ (2006) The effects of non-cas graphing calculators on student achievement and attitude levels in mathematics: A meta-analysis. *Sch Sci Math* 106(1):16–26

- Gatys LA, Ecker AS, Bethge M (2016) Image style transfer using convolutional neural networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition
- Gafni O, Polyak A, Ashual O, Sheynin S, Parikh D, Taigman Y (2022) *Make-a-scene: Scene-based text-to-image generation with human priors*. In: Proceedings of the European Conference on Computer Vision
- Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y (2020) Generative adversarial networks. *Commun ACM* 63(11):139–144
- Haas C (2013) *Writing Technology: Studies on the Materiality of Literacy*. Routledge, New York and London
- Hsu Y-C, Ching Y-H (2023) Generative artificial intelligence in education, part one: the dynamic frontier. *TechTrends* 67(4):603–607
- Ha AYJ, Passananti J, Bhaskar R, Shan S, Southen R, Zheng H, Zhao BY (2024) Organic or diffused: Can we distinguish human art from AI-generated images? *arXiv preprint arXiv:2402.03214*
- Huang R (2019) *Educational Technology a Primer for the 21st Century*. Springer, Singapore
- Hitsuwari J, Ueda Y, Yun W, Nomura M (2023) Does human-ai collaboration lead to more creative art? aesthetic evaluation of human-made and AI-generated haiku poetry. *Comput Hum Behav* 139:107502
- Huang Y, Wu W, Wu P, Teng Y (2024) Investigating the impact of integrating prompting strategies in ChatGPT on students learning achievement and cognitive load. In: Proceedings of the International Conference on Innovative Technologies and Learning
- Hwang G-J, Xie H, Wah BW, Gašević D (2020) Vision, challenges, roles and research issues of artificial intelligence in education. *Comput Educ: Artif Intell* 1:100001
- Hwang G, Yang L, Wang S (2013) A concept map-embedded educational computer game for improving students' learning performance in natural science courses. *Comput Educ* 69:121–130
- Isola P, Zhu J-Y, Zhou T, Efros AA (2017) Image-to-image translation with conditional adversarial networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition
- Jiang HH, Brown L, Cheng J, Khan M, Gupta A, Workman D, Hanna A, Flowers J, Gebu T (2023) AI art and its impact on artists. In: Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society
- Jeon J, Lee S (2023) Large language models in education: A focus on the complementary relationship between human teachers and ChatGPT. *Educ Inf Technol* 28(12):15873–15892
- Jia N, Luo X, Fang Z, Liao C (2024) When and how artificial intelligence augments employee creativity. *Acad Manag J* 67(1):5–32
- Javaheri Pour I, Abbasi E, Kian M, Hasanpour M (2021) A comparative study of primary school visual arts curriculum in Australia, Canada, Iran and Ireland. *Iran J Comp Educ* 4(2):1097–1116
- Jin Y, Yoon J, Andrew Self J, Lee K (2024) AI as a catalyst for creativity: Exploring the use of a generative approach in fashion design for improving their inspiration. In: Proceedings of the DRS Biennial Conference, pp. 1–27
- Keller JM (1987) Strategies for stimulating the motivation to learn. *Perform Instr* 26(8):1–7
- Kumar P, Gupta V, Grover M (2024) Dual attention and channel transformer-based generative adversarial network for restoration of the damaged artwork. *Eng Appl Artif Intell* 128:107457
- Kwegyiriba A, Mensah RO, Ewusi E (2022) The use of audio-visual materials in teaching and learning process in Effia Junior High Schools. *Tech Soc Sci J* 31:106
- Kasneci E, Seifler K, Küchemann S, Bannert M, Dementieva D, Fischer F, Gasser U, Groh G, Günnemann S, Hüllermeier E et al. (2023) Chatgpt for good? on opportunities and challenges of large language models for education. *Learn Individ Differ* 103:102274
- Kim J, Yu S, Detrick R, Li N (2025) Exploring students' perspectives on generative ai-assisted academic writing. *Educ Inf Technol* 30(1):1265–1300
- Lee U, Han A, Lee J, Lee E, Kim J, Kim H, Lim C (2024) Prompt aloud!: Incorporating image-generative AI into STEAM class with learning analytics using prompt data. *Educ Inf Technol* 29(8):9575–9605
- Lyu M, Li W, Xie Y (2019) The influences of family background and structural factors on children's academic performances: A cross-country comparative study. *Chin J Sociol* 5(2):173–192
- Lawan AA, Muhammad BR, Tahir AM, Yarima KI, Zakari A, Abdullahi II AH, Hussaini A, Kademi HI, Danlami AA, Sani MA et al. (2023) Modified flipped learning as an approach to mitigate the adverse effects of generative artificial intelligence on education. *Educ J* 12(4):136–43
- Lambert C, Philp J, Nakamura S (2017) Learner-generated content and engagement in second language task performance. *Lang Teach Res* 21(6):665–680
- Luong T-T, Tran M-T (2024) Designing a gai-assisted pedagogical task for teaching negotiation skills: A design thinking approach. In: Proceedings of the International Conference on Advances in Education and Information Technology
- Lu Y, Yang X, Li X, Wang XE, Wang WY (2024) LlmScore: Unveiling the power of large language models in text-to-image synthesis evaluation. In: Proceedings of the Advances in Neural Information Processing Systems
- Mayer RE (2002) Multimedia learning. In: *Psychology of Learning and Motivation* vol. 41, pp. 85–139. Elsevier
- Mao J, Chen B, Liu JC (2024) Generative artificial intelligence in education and its implications for assessment. *TechTrends* 68(1):58–66
- Marrone R, Crompton D, Medeiros K (2024) How does narrow AI impact human creativity? *Creativity Research Journal* (0), 1–11
- Mishra P, Koehler MJ (2006) Technological pedagogical content knowledge: A framework for teacher knowledge. *Teach Coll Rec* 108(6):1017–1054
- Margolis H, McCabe PP (2003) Self-efficacy: A key to improving the motivation of struggling learners. *Prev Sch Fail: Altern Educ Child Youth* 47(4):162–169
- Midgley C, Maehr ML, Hruda LZ, Anderman E, Anderman L, Freeman KE, Urdan T (2000) *Manual for the patterns of adaptive learning scales*. Ann Arbor: University of Michigan, 734–763
- Mordvintsev A, Olah C, Tyka M (2015) Deepdream-a code example for visualizing neural networks. *Google Research* 2(5)
- Mahajan S, Rahman T, Yi KM, Sigal L (2024) Prompting hard or hardly prompting: Prompt inversion for text-to-image diffusion models. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition
- Mulian H, Shlomov S, Limonad L, Noccaro A, Buscaglione S (2024) Mimicking the maestro: Exploring the efficacy of a virtual ai teacher in fine motor skill acquisition. In: Proceedings of the AAAI Conference on Artificial Intelligence
- Nichol AQ, Dhariwal P, Ramesh A, Shyam P, Mishkin P, McGrew B, Sutskever I, Chen M (2022) Glide: Towards photorealistic image generation and editing with text-guided diffusion models. In: *Proceedings of the International Conference on Machine Learning* (2022)
- Ngo TTA (2023) The perception by university students of the use of ChatGPT in education. *Int J Emerg Technol Learn* 18(17):4
- Noy S, Zhang W (2023) Experimental evidence on the productivity effects of generative artificial intelligence. *Science* 381(6654):187–192
- Papert SA (2020) *Mindstorms: Children, Computers, and Powerful Ideas*. Basic Books, New York
- Peppler KA (2010) Media arts: Arts education for a digital age. *Teach Coll Rec* 112(8):2118–2153
- Pavlik JV, Pavlik OM (2024) Art education and generative ai: An exploratory study in constructivist learning and visualization automation for the classroom. *Creative Educ* 15(4):601–616
- Paas F, Van Merriënboer JJ (2020) Cognitive-load theory: Methods to manage working memory load in the learning of complex tasks. *Curr Direct Psychol Sci* 29(4):394–398
- Reddy MDM, Basha MSM, Hari MMC, Penchalaiah MN (2021) Dall-e: Creating images from text. *UGC Care Group I J* 8(14):71–75
- Rombach R, Blattmann A, Lorenz D, Esser P, Ommer B (2022) High-resolution image synthesis with latent diffusion models. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition
- Ramesh A, Dhariwal P, Nichol A, Chu C, Chen M (2022) Hierarchical text-conditional image generation with clip latents. *arXiv preprint arXiv:2204.06125* (2022)
- Renninger KA, Hidi S (2015) *The Power of Interest for Motivation and Engagement*. Routledge, New York
- Radford A, Kim JW, Hallacy C, Ramesh A, Goh G, Agarwal S, Sastry G, Askell A, Mishkin P, Clark J (2021) Learning transferable visual models from natural language supervision. In: Proceedings of the International Conference on Machine Learning
- Ramesh A, Pavlov M, Goh G, Gray S, Voss C, Radford A, Chen M, Sutskever I (2021) Zero-shot text-to-image generation. In: Proceedings of the International Conference on Machine Learning
- Ruiz-Rojas LI, Acosta-Vargas P, De-Moreta-Llovet J, Gonzalez-Rodriguez M (2023) Empowering education with generative artificial intelligence tools: Approach with an instructional design matrix. *Sustainability* 15(15):11524
- Ruskov M (2023) Grimm in Wonderland: Prompt engineering with MidJourney to illustrate fairytales. *arXiv preprint arXiv:2302.08961*
- Singhal K, Azizi S, Tu T, Mahdavi SS, Wei J, Chung HW, Scales N, Tanwani A, Cole-Lewis H, Pfohl S et al. (2023) Large language models encode clinical knowledge. *Nature* 620(7972):172–180
- Skjuve M, Følstad A, Brandtzaeg PB (2023) The user experience of ChatGPT: findings from a questionnaire study of early users. In: Proceedings of the International Conference on Conversational User Interfaces
- Staddon RV (2020) Bringing technology to the mature classroom: age differences in use and attitudes. *Int J Educ Technol High Educ* 17(1):11
- Sullivan G (2010) *Art Practice as Research: Inquiry in Visual Arts*. Sage, California
- Sweller J, Van Merriënboer JJ, Paas F (2019) Cognitive architecture and instructional design: 20 years later. *Educ Psychol Rev* 31:261–292
- Sweller J (1988) Cognitive load during problem solving: Effects on learning. *Cogn Sci* 12(2):257–285

- Sweller J (2011) Cognitive load theory. *Psychol Learn Motiv* 55:37–76
- Tan WR, Chan CS, Aguirre HE, Tanaka K (2017) Artgan: Artwork synthesis with conditional categorical GANs. In: *Proceedings of the IEEE International Conference on Image Processing*
- Tyler CW, Likova LT (2012) The role of the visual arts in enhancing the learning process. *Front Hum Neurosci* 6:8
- Uswatun K, Mariani S et al. (2020) Comparison between generative learning and discovery learning in improving written mathematical communication ability. *Int J Instr* 13(3):729–744
- Vygotsky LS (1978) *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press, London
- Wang Z, Bergin C, Bergin DA (2014) Measuring engagement in fourth to twelfth grade classrooms: The classroom engagement inventory. *Sch Psychol Q* 29(4):517
- Wang X, Liu X-Q (2023) Potential and limitations of ChatGPT and generative artificial intelligence in medical safety education. *World J Clin Cases* 11(32):7935
- Wu Y, Mou Y, Li Z, Xu K (2020) Investigating American and Chinese subjects' explicit and implicit perceptions of ai-generated artistic work. *Comput Hum Behav* 104:106186
- Walsh JN, O'Brien MP, Costin Y (2021) Investigating student engagement with intentional content: An exploratory study of instructional videos. *Int J Manag Educ* 19(2):100505
- Yu H, Guo Y (2023) Generative artificial intelligence empowers educational reform: current status, issues, and prospects. *Front Educ* 8:1183162
- Zhang B (2023) Preparing educators and students for ChatGPT and AI technology in higher education. *ResearchGate*
- Zhou E, Lee D (2024) Generative artificial intelligence, human creativity, and art. *PNAS nexus* 3(3):052
- Zhu J-Y, Park T, Isola P, Efros AA (2017) Unpaired image-to-image translation using cycle-consistent adversarial networks. In: *Proceedings of the IEEE International Conference on Computer Vision*
- Zylinska J (2020) *AI Art: Machine Visions and Warped Dreams*. Open Humanities Press, London

Acknowledgements

This study was supported by the National Natural Science Foundation of China under Grants 62407042 and 62277045; the Natural Science Foundation of Shandong Province under Grant ZR2024QF075; and the Humanities and Social Science Fund of the Ministry of Education of China under Grant 24YJC880004.

Author contributions

Cunling Bian and Weigang Lu: Conceptualization of this study, Development of the instructional framework, Methodology design, Identification of key research questions, Writing and editing. Xiaofang Wang: Data analysis, Writing and editing. Yingxue Huang and Shanzhou: Practical implementation, Insights on aesthetic appreciation. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Ethical approval

This study was reviewed and approved by the Bioethics Committee of the Ocean University of China (Approval No. OUC-HM-20231205) on December 5, 2023, prior to the commencement of the research. The approval scope encompassed all procedures involving human participants, including recruitment protocols, data collection methodologies, and analytical approaches. The research strictly adhered to the ethical principles outlined in the 1964 Helsinki Declaration and its subsequent amendments, as well as institutional guidelines for human participant protection. All experimental protocols aligned with the committee's requirements for risk mitigation and data confidentiality.

Informed consent

Written informed consent was obtained from all participants (or their legal guardians for minors under 18 years old) by the research team between January and March 2024. Consent forms explicitly detailed: 1. The study's objectives, procedures, potential risks (minimal/no physical or psychological harm), and societal benefits. 2. Assurance of anonymity for participants in all published materials. 3. Authorization for data usage strictly for academic purposes, including publication in anonymized form. 4. Voluntary participation and withdrawal rights. For underage participants, guardians received additional verbal explanations from trained researchers to confirm comprehension before signing. Signed forms were stored securely in password-protected institutional databases. As a non-interventional questionnaire-based study, all participants were explicitly informed of their anonymized data's role in advancing learning behavior research, with no foreseeable risks to personal or professional standing.

Additional information

Correspondence and requests for materials should be addressed to Cunling Bian or Weigang Lu.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025