Exploring the Usage of Generative AI for Group Project-Based Offline Art Courses in Elementary Schools

ZHIQING WANG, Sun Yat-sen University, China
HAOXIANG FAN, Sun Yat-sen University, China
SHIWEI WU, Sun Yat-sen University, China
QIAOYI CHEN, Sun Yat-sen University, China
YONGQI LIANG, Xiangzhou Experimental School of Zhuhai, China
ZHENHUI PENG*, Sun Yat-sen University, China

The integration of Generative Artificial Intelligence (GenAI) in K-6 project-based art courses presents both opportunities and challenges for enhancing creativity, engagement, and group collaboration. This study introduces a four-phase field study, involving in total two experienced K-6 art teachers and 132 students in eight offline course sessions, to investigate the usage and impact of GenAI. Specifically, based on findings in Phases 1 and 2, we developed *AskArt*, an interactive interface that combines DALL-E and GPT and is tailored to support elementary school students in their art projects, and deployed it in Phases 3 and 4. Our findings revealed the benefits of GenAI in providing background information, inspirations, and personalized guidance. However, challenges in query formulation for generating expected content were also observed. Moreover, students employed varied collaboration strategies, and teachers noted increased engagement alongside concerns regarding misuse and interface suitability. This study offers insights into the effective integration of GenAI in elementary education, presents *AskArt* as a practical tool, and provides recommendations for educators and researchers to enhance project-based learning with GenAI technologies.

CCS Concepts: • **Human-centered computing** → **Field studies**; **Collaborative interaction**.

Additional Key Words and Phrases: Generative AI, project-based learning, art education, elementary school

1 Introduction

Art courses in elementary schools or K-6 education [1] are popular around the world. In these courses, instructors implement a variety of teaching methods with activities (*e.g.,* group or individual projects, discussions, exhibitions) to foster children's creativity, critical thinking, self-identity, and social

---

*Corresponding author

[1]In this paper, we use K-6, primary schools, and elementary schools interchangeably, which refer to the initial stage of formal education and typically serve children from around ages 6 to 11 or 12.

Authors' Contact Information: Zhiqing Wang, wangzhq83@mail.sysu.edu.cn, Sun Yat-sen University, Zhuhai, China; Haoxiang Fan, pengzhh29@mail.sysu.edu.cn, Sun Yat-sen University, Zhuhai, China; Shiwei Wu, wushw28@mail2.sysu.edu.cn, Sun Yat-sen University, Zhuhai, China; Qiaoyi Chen, chenqy99@mail2.sysu.edu.cn, Sun Yat-sen University, Zhuhai, China; Yongqi Liang, shiyanxuexiaoqiqi@126.com, Xiangzhou Experimental School of Zhuhai, Zhuhai, China; Zhenhui Peng, pengzhh29@mail.sysu.edu.cn, Sun Yat-sen University, Zhuhai, China.

---

competencies [16, 17]. Project-based learning (PBL) is one such method that is extensively applied in art education [2, 32, 43, 63]. This approach engages students in real-world projects and often encourages them to discuss ideas and work together in peer groups. The effectiveness of PBL has been well-documented in fostering students' creativity, autonomy, collaboration, problem-solving skills, and reflective thinking in real-world contexts [10, 22, 24, 46].

From the teachers' perspectives, a typical PBL process in art course can include planning (*e.g.,* topic, question, deadlines), testing (*e.g.,* project progress, facilitation), and reflecting (*e.g.,* presentation, explanation) [22]. This paper focuses on the testing phase of the PLB process, in which students in a small group need to work together to finish a task given by the teachers. During this phase, students gather relevant information, brainstorm ideas, and implement solutions. However, elementary school students may struggle with these tasks due to limited knowledge, teamwork skills, and self-directed learning abilities [8, 37]. In offline classroom where there is normally one art teacher in one class, students can query the teacher for support when they face challenges during information seeking and solution ideation. Nevertheless, the teacher's support may not be in time, especially when managing multiple groups with varying needs and abilities [37, 57, 67].

Recent advances in Generative AIs (GenAIs) offer a promising alternative to offering in-situ information seeking and ideation support to students in their PBL art courses. For example, the text-to-image generation techniques like DALL-E enable the production of high-quality images from textual descriptions [15, 44] and have been used in supporting ideation in various design-related tasks, *e.g.,* facilitating creativity [11], explore design possibilities[42], and enhancing collaborative ideation[30]. The conversational agents powered by large language models (LLMs, *e.g.,* GPT) can answer any user query in nearly real-time, which could bridge knowledge gaps and promote creativity [5, 13, 40, 66]. Prior studies have underscored the role of GenAIs in various educational roles, such as adaptive learning, collaborative learning, and AI-assisted content creation [40, 48]. Despite the potentials, little research has investigated the usage of GenAIs to support students in information seeking and ideation in their group-based offline art courses in elementary schools. Applying GenAIs in this education scenario is unique against using them in higher education [65] or individual art learning activities [25]. For one thing, many students in elementary schools may find the GenAIs non-transparent regarding their capabilities and implications [54]. For another, it is not always feasible, or often not recommended for group projects in offline classrooms, to offer each elementary school student a device for accessing GenAI in art courses. Besides, it is unknown how teachers would perceive and collaborate with GenAI to facilitate students in the PBL art courses. Correspondingly, three research questions are under-explored:

- **RQ1**: How would elementary students use GenAIs, *i.e.,* the text-to-image generation techniques and large language models in our case, for information seeking and solution ideation in their group projects in offline art courses?
- **RQ2**: What interaction design with GenAIs would be helpful for supporting students in PBL art courses?
- **RQ3**: How would GenAIs affect the the process of the PBL art course from teachers' perspectives?

Answering these questions can contribute to the CSCW communities with empirical findings and implications of facilitating elementary students in group-based information-seeking and ideation activities with GenAIs.

To this end, we collaborated with two experienced K-6 art teachers in China to run a four-phase field study (Fig. 1) to explore the usage of GenAIs in a total of eight-course sessions. In Phase 1, we introduced DALL-E, a text-to-image model, to the **Class A** with 24 students who form three 8-person groups in their project-based art course. In four course sessions, whenever they wish,

students in Class A can use DALL-E in the Chatbox [2] interface via one computer in the classroom. In Phase 2, we additionally introduced GPT, an LLM-powered conversational agent, to Class A for a one-course session. Results of Phases 1 and 2 answer RQ1 by highlighting DALL-E's values in inspiring ideas and GPT's benefits in providing background information and implementation guidance. We also found that students use GPT and DALL-E in a sequential order in Phase 2 but often struggled with query formulation to get the expected content. These findings informed our design of *AskArt*, an interface that aims to answer RQ2 by integrating GPT and DALL-E with tailored features to support K-6 students in art courses. Specifically, compared to the traditional interface in ChatBox, *AskArt* provides: i) Project-Related Introduction of GenAI; ii) Audio-to-Text Integration; iii) Suggested Follow-Ups; iv) Select and Generate.

To evaluate *AskArt*, in Phase 3, we deployed it to class A for another course session to have qualitative comparisons between *AskArt* and Chatbox. In Phase 4, we further compared *AskArt* to a baseline condition without access to any computer devices via a co-designed one-course session with Teacher 2 for Class B and Class C, in which only three randomly selected groups in each class can access *AskArt*. Results from Phases 3 and 4 provide additional insights into RQ1 by demonstrating GenAI's role in supporting information seeking and ideation in PBL art courses. They also inform RQ2 by highlighting the benefits of *AskArt*'s "Suggested Follow-Ups" and "Select and Add Content in GPT as the prompts" to DALL-E for fostering elementary students' interaction with GenAIs. As for RQ3, students with *AskArt* in Phase 4 reported significantly higher satisfaction than those without *AskArt* in the PBL art course. They developed different strategies for integrating GenAIs into their group work, balancing between collective brainstorming and individual tasks. Across the four phases, the two teachers noted in the interviews that GenAIs increased students' engagement in the courses but concerned the misuse of GenAIs in the group-based art projects. Based on our findings, we discuss opportunities and implications for using GenAIs to help elementary students seek information and ideate solutions during group project-based courses.

Our contributions to the CSCW communities are three-fold. First, we offer first-hand findings on the usage of GenAIs in group project-based K-6 art courses in offline classrooms. The findings showcase the benefits, challenges, usage patterns, and impact of GenAIs in these courses. Second, we present an interactive interface *AskArt* that customizes the interaction with GenAI for K-6 students to facilitate information seeking and solution ideation in their group project-based art courses. Third, we offer actionable insights for educators and researchers to leverage GenAIs in elementary students' group project-based learning and more broadly, group-based information-seeking and ideation activities.

In the following sections, we first review related work on PBL, GenAIs for supporting art activities, and GenAIs for education. Then, we present the overview of our four-phases studies (section 3), followed by the details and findings in Phases 1 and 2 (section 4), the design of *AskArt* (section 5), and studies in Phases 3 and 4 (section 6). Lastly, we discuss implications of our work in section 7.

2 Related Work

2.1 Project-Based Learning in Art Education

Project-based learning (PBL) represents a shift from traditional teacher-centered instruction to a more student-focused approach, where learning occurs within a specific context, students engage actively, and knowledge is shared through social interactions [28]. This method requires students to address real-world problems through extended inquiry, fostering active and personalized learning experiences [10, 51]. PBL aims to provide students with a broad skill set that transcends traditional educational methods, enhancing the practical aspects of learning and boosting employability [3].
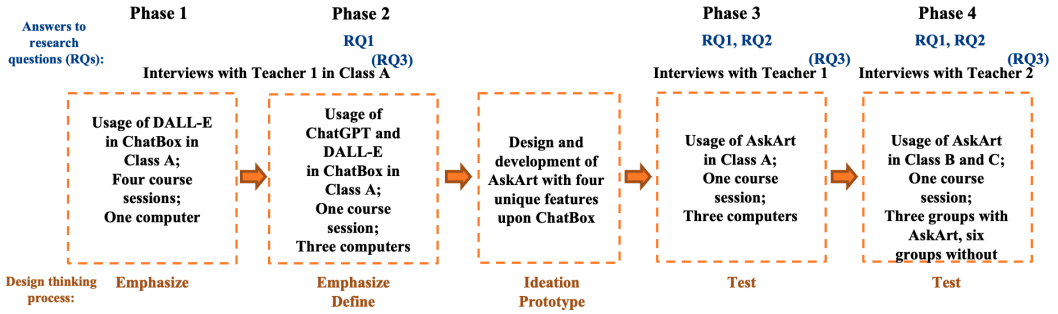
---

[2]web.chatboxai.app

Fig. 1. Overview of our four-phases study as well as its relationship to the design thinking process and three research questions, *i.e.,* RQ1 about elementary students' usage of GenAIs in their PBL art courses, RQ2 about helpful interaction design with GenAIs for supporting them, and RQ3 about teachers' thoughts on the impact of GenAIs on their courses.

The effectiveness of PBL has been validated in various practice [21, 37, 53, 59]. For example, a study in Slovakian secondary schools compared the outcomes of teaching economics through PBL with traditional methods [34]. It involved 123 students divided into control and experimental groups, with the latter participating in PBL. The findings showed that the experimental group performed better academically, confirming the effectiveness of PBL over traditional teaching methods.

PBL has also shown considerable promise in art education by engaging students in meaningful learning activities. In progressive education, particularly in the arts, emphasis is placed on hands-on activities and projects that encourage exploration and concept discovery through active involvement [41]. This aligns with the principles of PBL. Implementing PBL in art education has been confirmed to boost students' motivation, as students learn by doing and applying their ideas to complete directed projects, leading to more effective learning outcomes [36].

However, PBL in art education presents challenges from both student and teacher perspectives. Designing and preparing projects can be time-consuming for teachers [8, 67], and projects must be carefully designed to meet educational goals to keep students engaged [57]. Students often struggle to take the initiative in the learning process, where they must identify problems and apply their knowledge to find practical solutions [1, 4, 33]. Continuous guidance from teachers and facilitators is recommended to improve students' PBL skills, particularly in teamwork and self-directed learning [8]. However, this adds to the already substantial workload reported by teachers involved in PBL [37].

Our work is motivated by the benefits of PBL in art education and aims to mitigate these challenges in implementing group project-based art courses in elementary school students. Compared to the PBL practices mentioned above, our target scenario could be more challenging, as elementary school students usually are less experienced in independent project completion than senior students and the group PBL with two or more students would be more dynamic than the individual PBL.

2.2 Supporting Art Activities with GenAI

With the growing availability of GenAI technologies, including large language models (LLMs) and image generators, researchers are exploring opportunities for using GenAI to assist in art creation. Generative AI models have proven to be highly effective for art creation, particularly in image generation. For example, Ghosh et al. [18] introduced a GAN-based interactive image generation system designed to create complete images from partial user strokes or sketches. This

system also functions as a recommender, assisting users throughout their creative process to help them produce their desired images. With the introduction of DALL-E 3, MidJourney, and Stable Diffusion, high-quality AI image generation became available to the public and was actively involved in the creation process of artworks. For instance, Verheijden and Funk [56] introduced BrainFax, a tool that integrated with Stable Diffusion to help humans generate and edit images. Professional designers participated in the evaluation for BrainFax shared their enthusiasm and saw potential or use in their own work.

In addition to image generators, recent work has begun to explore how LLMs were involved in projects in terms of ideation [27, 52]. For example, Shaer et al. [47] explored the integration of LLMs into the creative process, focusing on both the divergence stage of idea generation and the convergence stage of idea evaluation and selection. They developed a collaborative group-AI Brainwriting ideation framework that incorporates an LLM to enhance the group ideation process. Their study evaluated the idea-generation process and the resulting solution space, demonstrating that integrating LLMs in Brainwriting can significantly enhance both the ideation process and its outcomes. The integration of multiple GenAI models can further offer comprehensive support for the art project process. For example, Sun et al. [50] utilized LLMs and Stable Diffusion to exemplify the comprehensive support of GenAI in facilitating art projects through ideation, problem-solving, and visual support. In their work, LLMs were used for generating narrative travelogues based on the location information, while Stable Diffusion was used to create the images of future aligning the travelogues and the photo of that specific location. This comprehensive approach illustrates how GenAI can support the entire artistic process, from ideation and storytelling to visual representation, thereby enhancing the overall creative workflow.

The successful usage of GenAI in art activities inspires us to leverage generative technology to support students in their group art projects. We contribute to this line of related work with first-hand findings and insights about how students perceive and interact with GenAI in the group project-based art courses.

2.3 GenAI for Education

Extensive research has explored the application of GenAI to support students' learning activities from various aspects [48]. For example, Kazemitabaar et al. [26] developed CodeAid, an LLM-powered programming assistant that provides timely and personalized feedback without revealing direct code solutions. CodeAid assists students by answering conceptual questions, generating pseudo-code with explanations, and annotating incorrect code with fix suggestions. Deployed in a 700-student programming class over a 12-week semester, CodeAid demonstrated how GenAI can offer immediate and customized guidance, which enhanced the learning experience by promoting cognitive engagement and providing valuable support tailored to individual student needs. Similarly, Khan Academy has begun deploying "Khanmigo" [7], based on GPT, to offer one-on-one tutoring for learners of all backgrounds and skill levels and to provide guided lesson planning for educators. In addition to personalized guidance, GenAI has been leveraged to support educational environments by fostering essential soft skills like teamwork and collaboration. For example, Breideband et al. [6] introduce the "CoBi" system, a multi-party AI partner designed to enhance students' collaborative skills by focusing on the relationship dimension of collaboration. Unlike traditional AI-driven tutoring tools that emphasize one-on-one instruction and personalized learning, CoBi assists students in co-negotiating classroom agreements along four dimensions: respect, equity, community, and thinking. By employing advanced speech and language technologies, CoBi monitors small group collaborative discourse for evidence of these agreements and visualizes the findings to motivate reflection. These feedback visualizations enable students to understand their strengths in

developing collaborative relationships, identify areas for improvement, and cultivate critical AI literacy skills.

While GenAI tools are more accessible to higher education students, lower education students remain an important group of users that cannot be neglected [55]. The UNICEF (United Nations Children's Fund) has developed guidelines for "AI and child rights policy" with international experts, emphasizing that "the rights of children, as current users of AI-enabled systems and the future inhabitants of a more AI-saturated world, must be a critical consideration in AI development" [35]. Much HCI research has been devoted to studying the integration of GenAI in lower education settings (*e.g.,* junior or senior high school). However, AI in elementary school education has received relatively little attention. Zhou et al. [65] organized past findings of integrating GenAI into lower education and proposed a conceptual framework to support researchers, designers, and educators in creating effective AI learning experiences for K-12 students. Van Mechelen et al. [54] summarize that despite various efforts, the education of emerging technologies, including AI, primarily focuses on middle school students, with few records specifically addressing elementary students. Although limited, attempts to integrate AI in elementary school education have proven effective and beneficial. For example, ChatScratch [9], which uses Stable Diffusion with ControlNet and LLMs, helps with project planning, asset creation, and programming. This tool promotes autonomous learning of Scratch for elementary school children.

However, among these few attempts, none have integrated GenAI into existing curricula for elementary school students within a real classroom setting. Much less is known about how elementary school children perceive and use GenAI in real practice and how teachers react to the integration of GenAI in a real classroom setting. Answering these questions is the aim of our work and can shed lights into equipping offline classrooms with GenAI.

3 Overview of Four-Phase Field Study

To explore the elementary students' usage of Generative AIs (GenAIs) (RQ1), identify rooms for improvements of GenAIs' interface (RQ2), and examine the impacts of GenAIs (RQ3) in group project-based art courses, we collaborate with two elementary school teachers (noted as Teacher 1 and Teacher 2) in a local city in China to conduct a four-phase field study in their offline classrooms (Fig. 1). While the settings of authentic project-based art classroom may make it difficult for conducting a rigorous controlled experiment due to noisy factors like students' background and teaching schedules, such settings can help us ensure the ecological validity of our study and gain insightful qualitative results. No alterations were made to the class environment or curriculum content, aside from the inclusion of the GenAIs interface. Given the autonomous nature of PBL classes, and in consultation with the teacher, students' usage of the GenAIs interface in the course was entirely voluntary. While students were encouraged to engage with GenAIs, no compulsion was applied.

We followed the design thinking process to plan the four-phase study and adjusted it based on students' and teachers' feedback. In the "Empathize" and "Define" stages, we conducted interviews with Teacher 1 to understand the practices and challenges of enacting project-based learning (PBL) in the art courses and identify what and how appropriate GenAI techniques can help. Specifically, Phase 1 serves as a warm-up phase that helps the teacher and students learn about the usage of GenAIs. As both Teacher 1 and our research team agreed that the image generator could support students in ideating and picturing the project outcomes in the art course, we started with **DALL-E** [3], which "understands significantly more nuance and detail than the previous DALL-E systems and allows users to easily translate their ideas into exceptionally accurate images". Phase 1 lasts for four course sessions, after which the research team and Teacher 1 met and discussed the learned

---

[3]https://openai.com/index/dall-e-3/

insights. Followed Teacher 1's suggestions, in Phase 2, apart from DALL-E, we allowed students to use GPT-4 Turbo, to examine whether and how it could help students seeking information and guidance related to the project. In Phases 1 and 2, students in each group can access the GenAIs via a traditional conversational interface **Chatbox** [4] (Figure 2). In the GPT-4 window (Figure 2 A), students can start with an initial prompt introducing the context (A1), pose questions (A2), and receive detailed text responses (A3). In the DALL-E window (Figure 2 B), students can submit image queries (B1) and view the generated images (B2). With the design goals derived from the findings of Phases 1 and 2, in the "Ideation" stage of the design thinking process, our research team and Teacher 1 brainstormed interaction features of an interface upon Chatbox that might improve students' efficiency in using GenAIs in their art projects. In the "Prototype" stage, we developed the *AskArt* prototype. Lastly, in the "Test" stage, we conducted one user study in Phase 3 and another study in Phase 4 to evaluate *AskArt*.

Fig. 2. The Chatbox interfaces used in Phases 1 and 2. The left portion of the figure illustrates the question-asking interface and the right one illustrates the image-generation interface.

Table 1 summarizes the main setups of the four-phase study. Specifically, Teacher 1 has 10 years of art education experience and two years of experience in applying group project-based learning (PBL) in her class. We conduct Phases 1, 2, and 3 in Teacher 1's class (Class A) with 24 students (4 males and 20 females, Grade 3 - 5). Students in this class were divided into three 8-person groups during the project-based art course. Teacher 2 has 20 years of art education experience and 8 years of experience in applying group PBL in her art courses. We conducted Phase 4 in the two classes taught by Teacher 2: Class B with 54 students (Grade 6), and Class C with 54 students (Grade 6).

4 Phases 1 and 2: Usage of DALL-E and GPT via a Traditional Conversational Interface

4.1 Procedure, Project Task, and Collected Data

Prior to the first course session in **Phase 1**, the research team interviewed Teacher 1, discussing the most appropriate GenAI techniques that can be introduced to the art course, determining the course sessions in which students can access the GenAIs, and deciding how students can interact with the GenAIs. As described in section 3, we agreed that students can use DALL-E in the Chatbox

---

Table 1. Overview of the settings, respondents to the questionnaire, and usage GenAI interface in our four-phases study in the art course sessions.

Phase	Participants	# Sessions	Questionnaire		GenAI	
			No.	Respondents	Interface	Models
1	Teacher 1	4		None	Chatbox	DALL-E 3
2	with Class A: three 8-students	1	2	P2-1 to P2-6		DALL-E 3, GPT-4
3	groups	1	3	P3-1 to P3-5		
4	Teacher 2 with Class B/C: three/five 6-students groups with/without AskArt	1 in Class B. 1 in Class C	4	P4-1 to P4-13 with AskArt	AskArt	DALL-E 3, GPT-4o
			4-all	P4-1 to P4-108		

interface in Phase 1. Teacher 1 suggested that we could allow students to access GenAIs in the course sessions for ideation and implementation of an artwork for the main project in that semester. The students' task in this project is designing and creating a handcrafted artwork representing one of the dragons from a Chinese tale "The Nine Sons of the Dragon", whose names are Qiu Niu, Ya Zi, Chao Feng, Pu Lao, Suan Ni, Bi Xi, Bi An, Fu Xi, and Chi Wen. Each of the three groups are assigned with three sons of the dragon to be considered in their projects. The task prompt for each group is: "Please design and make a handcraft to show [Name of the Dragon's Son]. Think about its special features and what it means in the ancient stories and Chinese culture."

In the first session of Phase 1, which is the Week 3 of the 2024 Spring semester, Teacher 1 introduced DALL-E to students through a short tutorial prepared by the research team, which shows several example use cases of DALL-E. As they wish, students can interact with DALL-E in Chatbox deployed on one single computer in the classroom, which is a typical device setting in the elementary schools in main cities in China . Phase 1 lasted for four continuous course sessions. At the end of the last session, the research team and Teacher 1 discussed insights and lessons learned from the four sessions and the plans for the coming course sessions. We collected audio recordings of the interviews with Teacher 1.

In **Phase 2,** apart from DALL-E, we also allow students to use GPT-4 Turbo in a separate chat window of Chatbox, as Teacher 1 suggests at the end of Phase 1 that a large language model may help students collect information and search guidance for completing a project. Phase 2 lasts for one 40-minute course session with the same project theme in Phase 1. Before the course session, we first interviewed Teacher 1, introducing the usage of GPT to Teacher 1 and discussing the coordination of the research team to run the course session with GPT and DALL-E. Then, three members of our research team came to the classroom. They first introduced the GPT to students with its example use cases. Each of the three groups is assigned one laptop and one member of our team as the teaching assistant (TA). The TA was only responsible for typing the prompts into GPT or DALL-E if necessary but did not directly participate in the group project. At the end of the course session, we distributed Questionnaire 2 to six randomly selected students, two from each group, and conducted quick interviews with them to make sense of their responses to the questionnaires. Questionnaire 2 asks: i) the needs and challenges during the information-seeking process before using GenAIs (*i.e.,* GPT and DALL-E) in this course session; ii) ratings for GenAIs' assistance in the ideation process (1 - strongly disagree, 7 - strongly agree); and iii) benefits and challenges of using GenAIs. After the course session, we also interviewed Teacher 1 about the students' perceptions of GPT and DALL-E, the impact of GenAIs on her class so far, the challenges encountered, and the plan for future course

sessions. Data collected in Phase 2 include: audio recordings of the interviews with Teacher 1 and students, students' prompts to GPT and those for DALL-E in the course session, responses of six students on Questionnaire 2, and screen recordings of the GenAIs' usage in the computers (with the audio input and camera on) for the three student groups.

4.2 Analysis and Results in Phases 1 and 2

4.2.1 Data Analysis. **Interviews**. For the audio recordings of interviews with teachers and students collected in Phases 1 and 2, we employed the reflex thematic analysis approach. Two researchers independently reviewed the transcripts and responses to identify key themes. These themes were synthesized into affinity diagrams and discussed with all authors to ensure consistency and accuracy. The outstanding themes from the interview with teachers answer the **RQ3** and are presented later in subsubsection 6.2.4, while those from the interview with students help to answer **RQ1** and are detailed in this subsection.

Students' Responses to Questionnaires. The collected students' responses to the questionnaires distributed in Phase 2 contribute to **RQ1**. To ensure the authenticity of the responses, we took several approach and filter unsuitable responses if applicable. For Questionnaire 2, if a student's response indicated a possible misunderstanding of the question, we reconfirmed their response during follow-up interviews. The qualitative responses (*i.e.,* students' comments) to the questions in Questionnaires 2 are analyzed via a reflex thematic analysis similar to that for the interview data mentioned above.

Prompts to GPT and DALL-E and Screen Recordings. The logged prompts used by students and the screen course recordings in Phase 2 are used to identify patterns and understand students' engagement with the tools, contributing to RQ1. Usage analytics were performed on the logged prompts. Content analysis was conducted to examine the screen recordings, coding the behaviors and interactions observed in the videos. Specifically, students' responses to the answers generated by GenAIs in Phase 2 were coded by two authors and classified into three categories: (1) perceived as correct and helpful, (2) perceived as correct but not helpful, and (3) perceived as incorrect.

4.2.2 Findings in Phases 1 and 2 for RQ1. The data analysis revealed students' behavioral patterns, perceived benefits, and encountered challenges when interacting with GenAIs in their group projects, which help to answer our RQ1 posted in the Introduction (Fig. 3), *i.e.,* "**RQ1**: How would elementary students use GenAIs, *i.e.,* the text-to-image generation techniques and large language models in our case, for information seeking and solution ideation in their group projects in offline art courses?".

Finding 1: DALL-E and GPT inspired students' project ideas and guided the idea implementation, and GPT additionally provided background information about the project. In Phase 2, we asked the six respondents to rate the level of their agreement with the statement "The generated content helps our group come up with more ideas" and got an average score of $5.67 (SD = 0.82)$, indicating that the GenAIs was helpful for solution ideation in the elementary students' art projects; 1 - strongly disagree, 7 - strongly agree. For example, during Phase 2, as observed by the TA, students in Group 2 initially had no idea about the design of the dragon's eyes. After prompting DALL-E with "images of Qiu Niu from the Nine Sons of the Dragon", they were inspired by the blue eyes of the dragon in the generated images. Besides, four out of six respondents to Questionnaire 2 in Phase 2 inform their challenges and needs for GenAIs' guidance in implementing their project ideas. For example, P2-5 stated the difficulties of designing the scales of the dragon, "*we could not decide the suitable color for the scales*". P2-4 noted that they "*did not know what materials to use for building a physical demo of the dragon*". The GenAIs can provide personalized guidance to address these challenges of idea implementation. For example, in Phase 2, the TA observed that students in Group 2 used GPT to figure out "*how to reinforce the body of the*
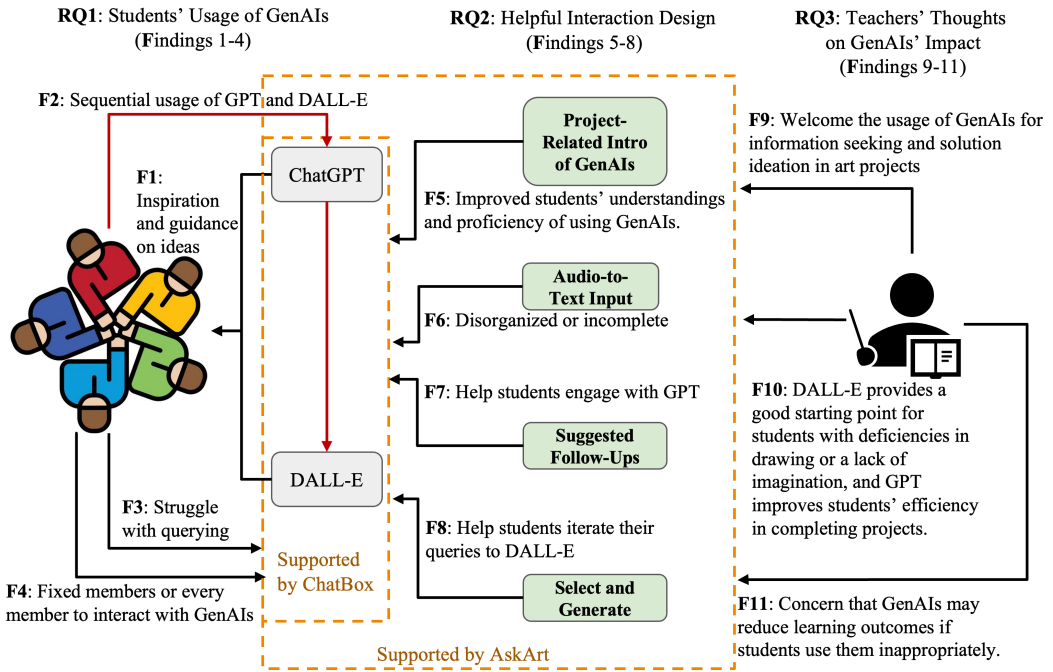
Fig. 3. Summary of our study **F**indings about elementary students' usage of GenAIs (F1-4 for RQ1), helpful interaction design (F5-8 for RQ2) explored via *AskArt*, and teachers' thoughts on GenAIs' impact (F9-11 for RQ3).

handcraft". P2-2 commented that the GenAIs "*can provide the most suitable solution based on our actual situations*". Additionally, three respondents mentioned that GPT was helpful in providing background knowledge and context about their projects. For instance, P2-6 noted that GPT helped understand the "*characteristics of the dragons*".

Finding 2: Students use GPT for information seeking and DALL-E for visual inspiration sequentially. In Phase 2, we introduced GPT to the students. By analyzing the interaction logs and the accompanied of screen recordings, we found that out of the total 28 queries to GPT and DALL-E, 13 queries were posed in a two-step process. In the first step, students queried GPT for information seeking. For example, Group 1 asked, "We want to make [the dragon named] Fu Xi. How can we use newspaper to make its head? We find this part very difficult". Group 3 asked for information about the theme of the project, "What does Bi Xi look like? He is one of the Nine Sons of the Dragon". Based on the information gathered from GPT, students then queried DALL-E for visual inspiration in the second step. For instance, Group 1 copied the complete answer from GPT and pasted it into the chat with DALL-E. Similarly, Group 3 asked DALL-E, "Images of Bi Xi from the Nine Sons of the Dragon" after getting a textual description of it from GPT.

Finding 3: Students struggled with querying, from typing issues to uncertainty in what to ask, leading to unsatisfying GenAI responses. Based on the TAs' observations, students frequently face difficulties in typing down their queries to GPT and DALL-E, which is a default input approach in Chatbox in Phase 2. The TAs in Groups 1 and 2 recalled that they were requested to help with typing over 90% of the time. Meanwhile, students were unsure of what questions were appropriate to ask to GPT and how to refine the queries, as P2-4 said, "*I did not know what*

questions I could ask [to GPT] at that moment. I also did not know I could follow up; I thought that was it". Consequently, students expressed dissatisfaction with the results returned by GenAIs, as indicated by three respondents in Phase 2. This dissatisfaction was further evidenced by 9 and 11 queries generating responses that were "perceived as incorrect" and "perceived as correct but unhelpful". For example, Group 1 asked for images about handcrafting with questions (*e.g.,* "How to make the back of [the dragon] Fu Xi from the Nine Sons of the Dragon?") in two out of their five queries to DALL-E, both of which were perceived incorrect by students. Group 2 asked GPT rather than DALL-E for "the face of [the dragon] Qiu Niu". Group 2 also inquired GPT about "the color of [the dragon] Qiu Niu's pupils" and, upon receiving an unsatisfactory answer, terminated the conversation and switched to using DALL-E. Group 3 requested "an image of [the dragon] Bi An" from DALL-E without providing any description and expected to get background information of Bi An, as recalled by the TA in Group 3. Similarly, they asked GPT to provide "an image of [the dragon] Suan Ni's handcraft", which misinterpret GPT's abilities as an image generator.

5 *AskArt*: Customizing Interactions with GenAIs for Facilitating Elementary Students in Art Projects

The findings in Phases 1 and 2 revealed the benefits of GenAIs for facilitating elementary students' information seeking and solution ideation in group-based art projects (Finding 1). However, students reported challenges in typing and formulating queries to GenAIs (Finding 3). Together with the students' patterns of interacting with GPT and DALL-E (Finding 2), our research team compiled three design goals and came up with interaction features with GenAIs for supporting students in PBL art courses. Our team consists of four master students majored in human-computer interaction (HCI) and one assistant professor specialized in HCI, and for one month, they met weekly to iteratively discuss, refine, and implement these interaction features in an interface we called *AskArt*. The three design goals of *AskArt* are as below.

DG1: *AskArt* should simplify students' input and tailor the initial prompts to GenAI based on the theme of the projects. This could reduce the complexity of typing prompts for elementary school students and improve the relevance of generated content to the art projects.

DG2: *AskArt* should support a seamless transition from the content generated by a text-based AI model to that produced by the image generator. This involves using GPT for text-based information gathering and then utilizing DALL-E for visual creation based on the gathered information. This transition between the usage of GPT and DALL-E should be intuitive and smooth.

DG3: *AskArt* should provide clear and effective guidance for prompting the GenAI models. Clear guidance can help users correctly understand and use GPT and DALL-E. This includes example use cases, templates, and adaptively suggested prompts that highlight the unique capabilities and best practices for each model.

5.1 *AskArt* Design and Implementation

Figure 4 and Figure 5 show the screenshots of the *AskArt*'s interface. The interface is divided into two main sections: "*Think Tank*" powered by GPT for text generation (Figure 4 left section), and "*little Painter*" powered by DALL-E for text-to-image generation (Figure 4 right section). In each section, users can scroll the chat window up and down. *AskArt* has the following unique features upon Chatbox based on the design goals.

Project-Related Introduction of GenAI. At the cold start, both "*Think Tank*" and "*Little Painter*" will send out introductory messages(Figure 4 A3 B1). These messages introduce the basic functionalities and user cases tailored to the PBL theme and tasks (**DG3**). The "*Think Tank*" interface also features two initial prompts in the Follow-Up Questions section, designed to help students
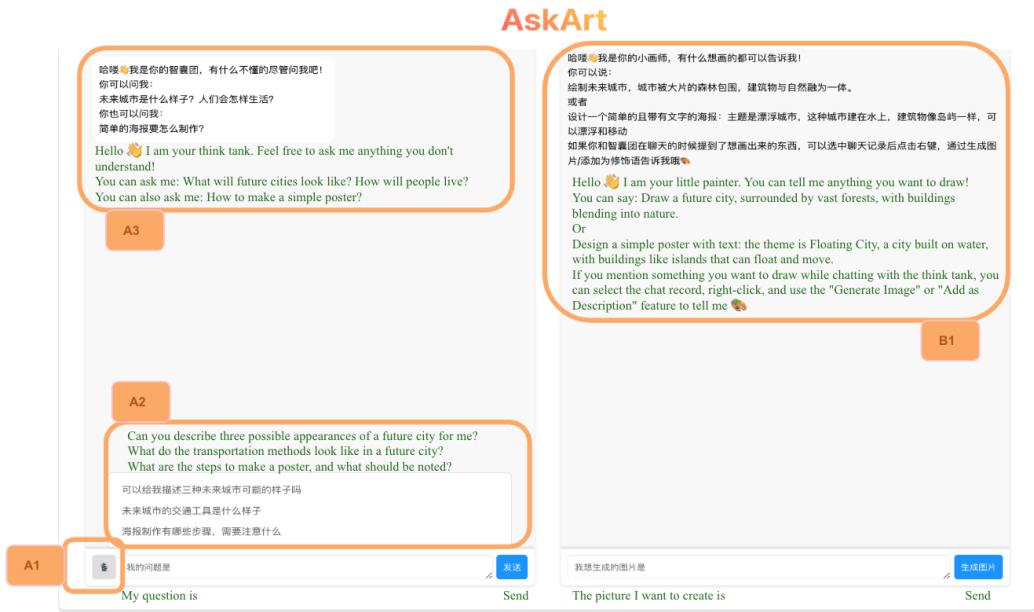
Fig. 4. *AskArt* Interface for Coldstart. The left part is the chat interface with "*Think Tank*" (powered by GPT), featuring: Audio-to-Text input (A1), Suggested Follow-Ups (A2), and Project-Related Introduction of "*Think Tank*" (A3). The right part is the chat interface with "*Little Painter* "(powered by DALL-E), featuring: Project-Related Introduction of "*Little Painter*" (B1).

understand its role and provide starting points for interaction (Figure 4 A2). These introductions and initial prompts were carefully crafted by our research team and teachers, and they were manually inputted before Phases 3 and 4 based on the corresponding project themes and tasks.

Audio-to-Text Input. We implemented an audio-to-text input feature in the "*Think Tank*" section (Figure 4 A1). This feature allows students to vocalize their queries, which are then converted to text using Web Speech API. This feature could ease the students' input effort to GPT (**DG1**).

Suggested Follow-Ups. To encourage students to seek further information or clarification, we introduced "*Suggested Follow-Ups*" at the top of the GPT input bar (**DG3**). These suggestions are generated based on the current conversation and update dynamically as the dialogue progresses. This feature aims to make students aware that they can provide feedback or refine their requests, and it also brings new angles for inquiry that students may not have considered, which could foster creativity and thoroughness.

Select and Generate. Another key feature in *AskArt* is the ability for students to select sentences or words from the "*Think Tank*" chat to use as prompts or parts of the prompt for "*Little Painter*" (**DG2**). To help students use this feature, each GPT response in "*Think Tank*" interface ends with an automatically generated draft prompt for DALL-E in "*Little Painter*" interface (Figure 4 A4). This prompt is based on the content of the corresponding GPT response, allowing students to directly use it if they wish to visualize the information provided by GPT. This not only aligns with the user pattern of sequentially using GPT and DALL-E (Finding 2) but also dynamically demonstrates to students how to construct effective queries for DALL-E (Finding 3). Offering a ready-to-use prompt could also simplify the input process (**DG1**) and encourage students to visualize their text-based
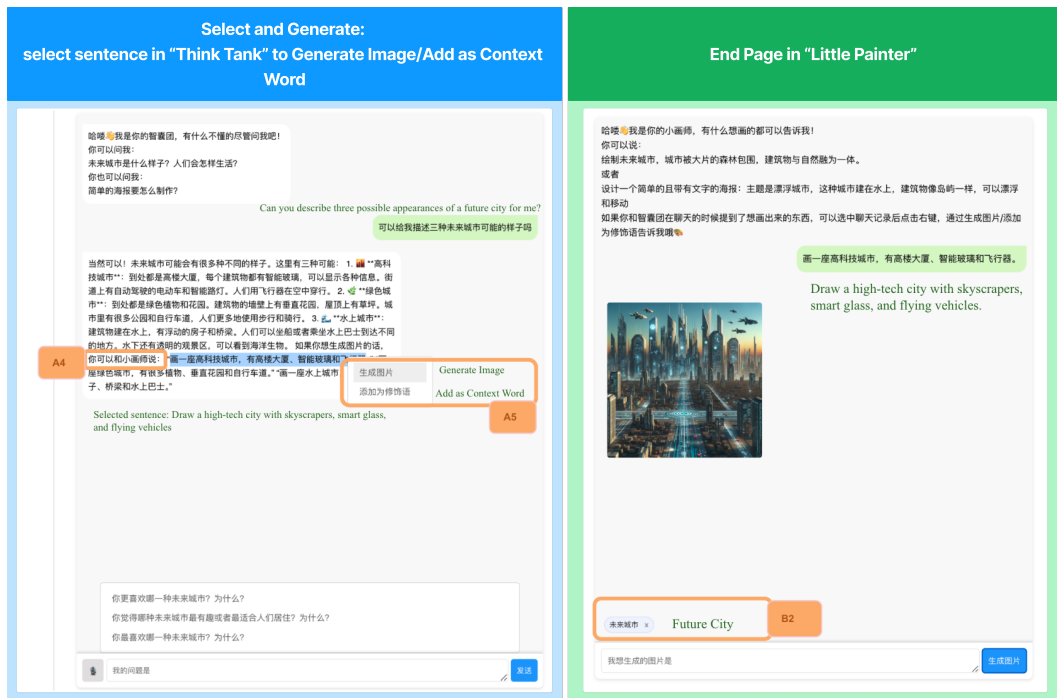
Fig. 5. *AskArt* Interface for *Auto-Generated DALL-E Prompts* and *Select and Generate*. The blue section is the interface of "*Think Tank*", showing the selection of *Auto-Generated DALL-E Prompts* (A4) followed by a right-click to use *Select and Generate* (A5). The green section is the end page in the interface of "*Little Painter*", demonstrating the resulting image sent by "*Little Painter*" after the actions in the blue section. B2 represents the *Context Word*.

findings immediately (Figure 5 A4). After selecting the desired text, students can right-click to access a menu (Figure 5 A5) with two options:

- *Generate Image*: Selected content is copied directly into DALL-E's input box, indicating the main subject for visualization.
- *Add as Context Word*: Selected content appears above DALL-E's input box and can be manually deleted if needed. This ensures that essential thematic or stylistic elements are included in the visualization, preventing omissions that might lead to unsatisfactory results. Context words can include PBL theme information or specific elements about art, such as colors and styles that students want to incorporate into their artwork.

This structured approach can prevent students from overloading the input box in DALL-E with all information, clearly distinguishing between the main subject of the image and additional contextual details. This feature can also encourage students to include relevant background information and stylistic preferences, which would improve the accuracy of DALL-E's visual output (**DG3**).

AskArt is developed using the Vue.js framework for the front end, and Python with Flask for the back end. Data is stored locally on the website for improved performance and quick access. The system processes user inquiries related to art by sending requests from the Vue.js interface to the Flask server, which then interacts with OpenAI's GPT-4o and DALL-E 3 API to generate responses. These responses are processed and returned to the user through the Vue.js front end. The project

ensures that all user interactions and queries are recorded and utilizes real-time data to provide accurate information.

Fig. 6. The samples of students' interactions with *AskArt* and the corresponding outcome. (a) Group 2 in Phase 3. (b) Group 1 in Phase 4.

6 Phases 3 and 4: Evaluation of *AskArt*

After developing *AskArt*, we collaborated with Teacher 1's teaching team and evaluated *AskArt* in their offline art courses.

6.1 Procedure, Project Task, and Collected Data

Phase 3 happened in Week 14 of 2024 Spring semester. In the interview with Teacher 1 before the course session, we introduced *AskArt* to the teacher and gathered her initial feedback. Different from Phases 1 and 2, as the semester progressed, the task for each student group changed to "using dragon-related text and images to create a poster that blesses the high-school students in 'Gao Kao', the college entrance exam in China" (Fig. 6(a)). Meanwhile, similar to Phases 1 and 2, this task requires students to work in group to ideate and implement an artwork (poster, specifically) as the outcome, during which we can examine how GenAIs would help. In the course session, similar to Phase 2, three members of the research team acted as TAs in three student groups, each with a laptop that have access to *AskArt*. At the end of the course session, we distributed Questionnaire 3 to six randomly selected students, two from each group, and collected five responses, as one student was not available to complete the questionnaire at that time. Questionnaire 3 asks: i) challenges encountered during class; ii) benefits and challenges of using *AskArt*; iii) ratings for GenAIs' assistance in information seeking and implementation guidance (1 - strongly disagree, 7 - strongly agree); iv) ratings for the assistance provided by *AskArt*'s specific features: *Audio-to-Text Input*, *Suggested Follow-Ups*, *Select and Generate* (1 - strongly disagree, 7 - strongly agree). We conducted quick interviews with the five respondents to make sense of their responses to the

questionnaires. We also interviewed Teacher 1 after the course session. We discussed students' perceptions of GPT, DALL-E, and *AskArt*, assessed the impact of GenAI in these sessions so far, and discussed potential ways to improve *AskArt*. The **data collected in Phase 3** include: audio recordings of the interviews with Teacher 1 and students; Students' prompts to GPT and DALL-E; responses of five students on Questionnaire 3; and screen recordings of *AskArt*'s usage in the computers (with the audio input and camera on) for the three student groups.

Phase 3 could help us qualitatively examine the specific features of *AskArt* against Chatbox, as the students in Class A used both tools in group-based projects with a similar solution ideation process. However, it can not enable us to evaluate *AskArt*'s impact on the students' PBL process compared to the cases without *AskArt*. **Phase 4** sought to address this limitation with a between-subjects study design. Specifically, different from those in Phases 1 to 3, students in Phase 4 were all from grade 6 and were assigned to Class B and Class C, both with 54 active students (29 males and 25 females). Each class was divided into 8 groups, each with 6-8 students. We collaborated with Teacher 2, who was responsible for the art courses of Classes B and C, to set up our study. In the interview with Teacher 2, before the course session, the research team introduced *AskArt* to the teacher and gathered initial feedback. They decided to conducted the study in week 16 of 2024 Spring semester, in which the course theme is "*Future City Poster*". In this course session, the task for each student group was to create posters depicting a future city (Fig. 6(b)). A total of eight groups were formed in each class (Class B and Class C), with six groups assigned to the experimental condition and ten groups to the baseline condition. For the experimental condition, each group used AskArt. For the baseline condition, no additional tools were provided, reflecting the standard classroom setting for normal PBL classes without GenAIs, as advised by the teachers. Teachers typically allow students to access tools like search engines at home but avoid providing such tools during class sessions due to concerns that access to search engines and social media platforms might distract students with unrelated content and hinder their focus.

At the end of the course session in Phase 4, we distributed Questionnaire 4-all to all students and collect 108 responses. **Questionnaire 4-all** asks for ratings on i) overall challenge in the class; ii) overall satisfaction; iii) participation in team collaboration; and iv) expression of creativity in class; 1 - the lowest score, 7 - the highest score. We also distributed Questionnaire 4 to the students in groups with *AskArt* and collect 13 responses. **Questionnaire 4** uses the same questions as Questionnaire 3, except for i) challenges encountered during class, to avoid repetition with Questionnaire 4-all. After the course session, we conducted quick interviews with in total of 3 respondents to make sense of their responses to the questionnaires. Lastly, we interviewed Teacher 2, discussing students' perceptions of GPT, DALL-E, and *AskArt*, assessing the impact of GenAI in this course session, and discussing potential ways to improve *AskArt*. The **data collected in Phase 4** include: audio recordings of the interviews with Teacher 2 and students; students' prompts to GPT and DALL-E; responses from all 108 students on Questionnaire 4-all and responses from 13 students from the groups with *AskArt* on Questionnaire 4; and screen recordings of *AskArt*'s usage in the computers (with the audio input and camera on) for the three student groups.

6.2 Analysis and Results of Phases 3 and 4

6.2.1 Data Analysis. Similar to subsubsection 4.2.1, we employed the reflex thematic analysis approach to analyze the audio recordings of interviews with teachers and students collected in Phases 3 and 4. The outstanding themes from the interview with teachers answer the **RQ3**, while those from the interview with students complement the Findings 1-3 of **RQ1** and answer **RQ2**.

For the questionnaire data, we first filtered unsuitable responses by carefully reviewing each response. For Questionnaire 4-all, we observed that some students left comments under the rating questions. If there was a mismatch between the attitudes revealed in the comment and the rating

score (*e.g.,* , a negative comment for a question rated 7 out of 7), these responses were filtered out, resulting in the exclusion of 9 out of 108 questionnaire sheets. For the students' ratings on the remained Questionnaire 4-all, we performed unpaired t-tests to statistically compare the differences between the groups (43 valid respondents) with and without (54 valid respondents) *AskArt* regarding students' perceived challenges, satisfaction, team collaboration, and creativity in the course session.

For the prompts to GPT and DALL-E and screen recordings, we used the similar analysis approach in subsubsection 4.2.1 to code the user behaviors and interactions with *AskArt*. Students' responses to the answers generated by GenAIs in *AskArt* were also classified into three categories: (1) perceived as correct and helpful, (2) perceived as correct but not helpful, and (3) perceived as incorrect.

(a)Distribution of Queries Generating Different Types of Student Responses
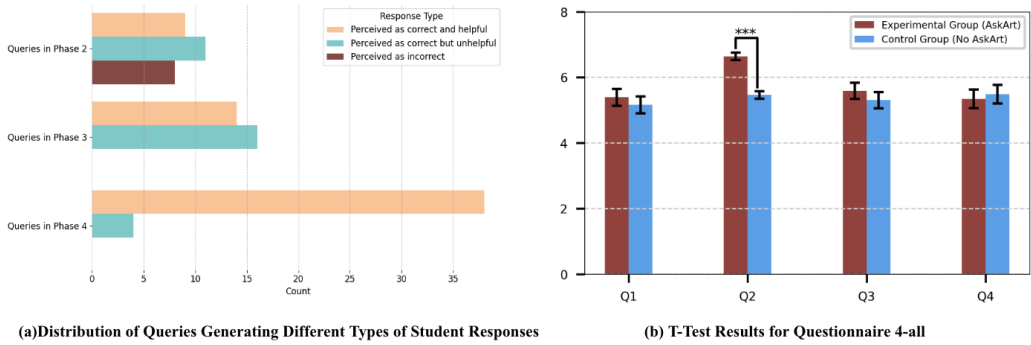
(b) T-Test Results for Questionnaire 4-all

Fig. 7. Results for Content Analysis for Students' Prompts in Phases 2-4 and t-test in Phase 4. (a) shows the responses of students to the answers generated by GenAI based on their queries, with bars representing the number of queries ending in three types of responses. (b) shows the t-Test results comparing n student ratings of their PBL class experiences between those who used AskArt (experimental group) and those who did not (control group). Q1 is "How challenging did you find this lesson?" (1: very low, 7: very high); Q2 is "Rate your satisfaction with this class."(1: very unsatisfied, 7: highly satisfied); Q3 is " actively participated in team collaboration during this lesson."(1: strongly disagree, 7: strongly agree); Q4 is "I fully expressed my creativity in this less. "(1: strongly disagree, 7: strongly agree). The results indicate a significant difference in overall satisfaction (Q2) between the groups, with the experimental group rating higher. Error bars represent standard deviations, with *** indicating a p-value < 0.001.

6.2.2 Findings in Phases 3 and 4 for RQ1 .

Students' responses to the questionnaires in Phases 3 and 4 complemented **Finding 1** (subsubsection 4.2.2) about GenAIs' helpfulness in inspiring project ideas, guiding idea implementation, and providing background information about the project. For example, in Phase 4, P4-13 commented the benefits from GenAIs as "it provides us with creative ideas for the project". The teaching assistant's (TA) observation of Group 3, where P4-13 belonged, along with this group's usage logs of *AskArt*, illustrates how GenAIs inspired their project. During the brainstorming session, they initially inquired with GPT about what a "Future City" might look like. They were particularly inspired by several elements mentioned by GPT, especially the idea of "hanging gardens". Subsequently, they used DALL-E to generate a visualization of a "Future City with hanging gardens" through the *Select and Generate* feature. The generated image further inspired them, leading to the inclusion of "hanging gardens" in their final poster. Besides, P3-2 in Phase 3 emphasized the effectiveness of GPT in providing background information about the theme of the project, particularly when he "*searched for blessings related to dragons*". P3-4 noted that he/she learned from GPT's answers about "*idioms related to dragons*". In Phase 4, P4-2 highlighted that the GPT enabled the group to "*quickly understand the project's theme about future city*".

In Questionnaire 4-all (Fig. 7), students using AskArt ($M = 6.64, SD = 0.12$) reported significantly higher ($t = 3.62, p = 0.0005$) overall satisfaction compared to those in the baseline group ($M = 5.46, SD = 0.27$). However, there was no significant difference ($t = 0.57, p = 0.57$) in perceived challenge level between the experimental group ($M = 5.39, SD = 0.26$) and the baseline group ($M = 5.16, SD = 0.28$). Similarly, for creativity expression, there was no significant difference ($t = -0.37, p = 0.72$), and the baseline group reported a slightly higher mean score ($M = 5.49, SD = 0.28$) compared to the experimental group ($M = 5.34, SD = 0.28$).

In Phase 4, students with *AskArt* in Classes B and C did not use GenAIs in their offline art courses before. We observed their interaction patterns with *AskArt* at a group level and got **Finding 4: Student groups either assigned fixed members or allow every member to interact with *AskArt* in the art courses**. For example, in Group 2, the emerging group leader assigned two students to interact with *AskArt*. These students were responsible for collecting questions from the rest of the group, formulating and posing queries to *AskArt*, and presenting the generated answers back to the group. Using this strategy, Group 2 still left some students, who were not engaged in group collaboration, uninvolved in the interaction with *AskArt*. In contrast, in Groups 6, any student with a question can interact with *AskArt*. This back-and-forth interaction with GenAIs allowed for individual adjustments based on the AI's suggestions. If the suggested adjustments impacted the whole design, the student would show the related GenAIs answer and discuss within the group; otherwise, they would make the decision independently and keep the GenAIs answer to themselves without sharing to the group. In this case, students, who were not engaged in group collaboration or did not pose any queries to *AskArt*, were uninvolved in the interaction with *AskArt*. Team engagement ratings (Fig. 7(b)) further support these observations, showing no significant difference ($t = 0.70, p = 0.48$) between the experimental group ($M = 5.59, SD = 0.24$) and the baseline group ($M = 5.31, SD = 0.30$).

Findings 2 and 3 (subsubsection 4.2.2) in Phases 1 and 2 informed the specific features of *AskArt*, and we found that these features streamlined the students' interactions with GPT and DALL-E. We described findings related to each feature in the next subsubsection.

6.2.3 Findings for RQ2 . Our evaluation studies of *AskArt* provided empirical understandings to **RQ2** (Fig. 3), *i.e.,* "What interaction design with GenAIs would be helpful for supporting students in PBL art courses?"

Finding 5: The "Project-Related Self-Introduction of GenAIs" improved students' understandings and proficiency of using GenAIs. Figure 7 (a) summarizes our analyses of the students' interaction logs with GenAIs. Specifically, we calculated the number of students' incorrect usages of GPT and DALL-E, represented by the number of queries generating "perceived as incorrect" responses (e.g., using DALL-E for instructional steps and GPT for image information). Our findings suggested that providing students with an introduction and user cases of GenAIs at the interface could improve their ability to use GPT and DALL-E correctly. This improvement is indicated by the disappearance of queries generating "perceived as incorrect" responses in Phases 3 and 4 (Figure 7 (a)). The introduction described GPT as a "think tank" and DALL-E as a "little artist", which helped students distinguish that GPT was not meant for visualization and DALL-E was not for textual information seeking. No instances of incorrect usage for GPT and DALL-E were observed in both Phase 3 in which students with prior experience using GenAIs in previous phases and Phase 4 in which students used GenAIs for the first time. For example, in Phase 3, students in Group 1 queried GPT for textual answers, *e.g.,* "Are there any idioms related to dragons?", and prompted DALL-E to visualize their ideas, *e.g.,* "A cool version of Ying Long", which demonstrates their improved understanding of the GenAIs' capabilities.

Finding 6: The "Audio-to-Text Input" to GPT was disorganized or incomplete, which was not helpful for supporting students in PBL art courses. In Phases 3 and 4, we introduced *AskArt* that allows students to query the GPT via voice input (Figure 4 A1). Students actively used this feature in 16 out of a total of 25 queries to GPT. However, we observed that in many cases, the voice input was disorganized or incomplete. For example, in Phase 3, students in Group 3 spoke "blessings related to dragons" to *AskArt*, which output many positive idioms that contain the Chinese word (*i.e.,* "Long") about dragon for New Year's (*e.g.,* Good luck in the Year of the Dragon – in Chinese: Long Nian Da Ji), birthday, and wedding blessings (*e.g.,* prosperity brought by the dragon and the phoenix – in Chinese: Long Feng Cheng Xiang). However, what these students needed were the blessings related to dragons for passing important exams. A complete query to GPT for this purpose would be "we are making a poster about passing important exams and need blessings related to dragons", which would output Chinese idioms like "Li Yue Long Men", wishing people, like the carp leaping over the dragon gate, achieve excellent results in the examinations and have a bright future. Similarly, in Phase 4, students in Group 6 talked to *AskArt*, "What words should be written in the painting of the Hanging Gardens?", while the organized and completed prompt, as observed by the TA, should be "I am making a poster of the Hanging Gardens. What slogan should be written on it?". One TA highlighted that this issue was particularly common when students were either nervous using the audio input feature for the first time or posed questions without organizing their thoughts beforehand. The TA explained "*Sometimes I needed to reconfirm the intended meanings of their voice input, which, of course, can not be understood by GPT*". These observations emphasize the ongoing need for improving the voice input to GenAIs (*e.g.,* via auto-completion of the transcribed text) to make them easier to use for elementary school students, which we will discuss in the Discussion section.

Finding 7: The "Suggested Follow-Ups" helped elementary students engage with GPT for seeking needed information. *AskArt* offers a follow-up question panel (Figure 4 A2) above the input bar. The respondents of Questionnaire 3 and Questionnaire 4 highly appreciated the helpfulness of this panel, with an average score of 6.06 (SD = 1.51); 1 - not helpful at all, 7 - very helpful. For example, in Phase 3, Group 2 utilized the follow-up questions in the "Suggested Follow-Ups" feature twice. First, they asked for more blessings containing the word "dragon", and then they requested additional examples that did not repeat the previous answers. This sequence of follow-ups concluded with the students perceiving the final response as correct and helpful (Figure 7 (a)). Additionally, students began independently providing feedback. We also observe that students in Group 1 provided feedback to the GPT's responses. They initially queried GPT with "Are there any idioms related to dragons for the college entrance exam?". When finding the response inadequate, they further queried GPT in the next turn with "I need more relevant examples". GPT has the ability to conduct such multi-turn conversation and provided Group 1 with more related idioms.

Finding 8: The "Select and Generate" feature helped students iterate their queries to DALL-E. *AskArt* enables users to select words or sentences from the chat with GPT and use them as *Context Word* for generating images in DALL-E (Figure 5). Respondents of Questionnaire 3 and Questionnaire 4 rated this feature highly helpful (M = 6.12, SD = 1.37); 1 - not helpful at all, 7 - very helpful. During Phase 3, students in Group 1 utilized this feature to make use of answers in GPT to form 5 out of 7 queries to DALL-E, and students in Group 3 utilized it for 5 out of 8 queries to DALL-E. An example of students using this feature in Phase 4 is shown in Figure 6 (b). In Phases 3 and 4, we observed that students iteratively adjusted the main prompts and *Context Word* (Figure 5 B) to DALL-E based on the images generated in the last turn. For example, in Phase 3, students in Group 1 iteratively refined their queries by making specific changes at each step. They began with "dragon's roar and phoenix's cry" without any *Context Word*. They then added "a handsome

version of Yinglong", incorporating "dragon's roar and phoenix's cry" as a *Context Word*. Next, they changed it to "a cute version of Ying Long with wings", adding "majestic dragon and phoenix's grace" along with the existing "dragon's roar and phoenix's cry" as *Context Word*. Finally, they settled on "a handsome version of Ying Long with wings", further adding "cloud dragon and wind tiger" to the list of *Context Word*. Similarly, Group 2 students started with the query "related to the [dragon] Zhuang Yuan". They refined this by changing it to "Zhuang Yuan in a red robe". Next, they modified it to "cartoon Zhuang Yuan in a red robe", and finally, they specified it further to "young cartoon Zhuang Yuan in a red robe" (Figure 6 (a)). This iterative process indicates that students can progressively enhance their queries by adding and adjusting *Context Word* to get needed images from DALL-E.

6.2.4   Findings for RQ3 . The interviews with Teacher 1 in Phases 1-3 and Teacher 2 in Phase 2 contributed understandings to RQ3 from the teachers' perspectives. We summarized the two teachers' perceived values, impact, and concerns of GenAIs on the students' PBL process after running the course sessions (Fig. 3).

Finding 9: Teachers welcomed the usage of GenAIs for information seeking and solution ideation in art projects. First of all, before running the course sessions, both teachers expressed strong support for and interest in integrating GenAIs interfaces into the students' project-based learning (PBL) process. Teacher 1 highlighted the relevance of GenAIs to PBL and art courses, "GenAIs are incredibly powerful and already widely used in professional contexts. It aligns well with the inquiry-based nature of PBL, and I am very optimistic about its integration into both group project-based art courses and other type of art courses in the future". Teacher 1 also noted that when interacting with GenAIs, especially in the *AskArt* interface, students were less likely to be distracted by irrelevant information compared to using web search engines for information seeking. "*In previous classes [without GenAIs], when we use computers to open web pages, we can not ensure that the pages only show information related to our courses, and students are easily distracted. GenAIs helped us mitigate this issue in the recent courses*".

Finding 10: DALL-E provided a good starting point for students with deficiencies in drawing or a lack of imagination, and GPT improved students' efficiency in completing projects. Teachers mentioned that with the assistance of GenAIs, students demonstrated an increasing engagement in completing the challenging projects. For students with deficiencies in drawing or a lack of imagination, DALL-E provided a crucial starting point by generating initial visuals. Teacher 1, who participated in Phases 1-3, noted that DALL-E provided a good starting point for students with deficiencies in drawing or a lack of imagination, "*Project-based courses can be difficult for students, and they often need step-by-step guidance on creating artworks. It can be challenging for students who are not skilled in drawing or designing to express their ideas in artwork. DALL-E helped them initiate the creative process and express themselves more freely in their artwork*". As for GPT, teachers highlighted its value in improving students' efficiency in completing projects. They remarked that students could search for all the necessary information within a single interface and receive tailored answers to their specific questions, thereby "*shortening the time spent on information gathering*". Teacher 1 pointed out in Phase 3, "*students needed to come up with words related to both dragons and the college entrance exam as well as finish the drawing of a poster within one 40-minute course session. Time was very limited, but with GenAIs, they succeeded*".

Finding 11: Teachers concern that GenAIs may reduce learning outcomes if students use them inappropriately. Teachers expressed concerns about the potential for GenAIs to undermine learning outcomes when students fail to use them appropriately. Teacher 2 emphasized the importance of helping students effectively incorporate GenAIs into their regular PBL activities, highlighting the need for both teachers and students to learn how to balance their time and effort

when interacting with these tools. She noted, "*both us and students need time to figure out how to effectively integrate GenAIs into their regular art courses and manage their time and effort interacting with GenAIs.*" Teachers also emphasized that the appropriateness of using GenAIs depends on the nature of the art project and the students' abilities. For example, they were more open to GenAI use in tasks requiring information seeking and solution ideation, such as handcraft projects in Phases 1 and 2 or poster designs in Phases 3 and 4. However, Teacher 1 stressed that students' skill levels should inform how GenAIs are used. She explained, "*If this group of students are not good at drawing and crafting objects, it is acceptable that they directly mimic the content generated based on their ideas, which can train their basic art skills. However, if the group is proficient in drawing and crafting objects, it is not acceptable that they copy the generated content in their artworks, as I would expect that they can enact critical and creative thinking to come up with unique artworks.*" Additionally, teachers expressed concern about projects that could be easily completed by directly using GenAI-generated content without further processing. For instance, if students relied solely on AI-generated materials to appreciate a painting or replicate an object, the use of GenAIs might bypass opportunities for learning. Teacher 1 suggested that students should submit the AI-generated materials they referenced along with their final artwork, explaining, "*I can assess whether they simply copy the materials.*"

7 Discussion

Based on our findings, this section discusses implications for the usage of GenAIs in group project-based learning activities, teachers' roles in GenAI-enhanced art courses, and future GenAI-powered learning support tools for elementary students.

7.1 Elementary Students' Use of Generative AIs in PBL

Based on the findings of RQ1 about elementary students' usage of GenAIs in a Chinese PBL art course, our study provides insights into the opportunities and challenges in these interactions. Firstly, our findings demonstrate that elementary school students can understand and effectively use GenAI tools. Despite their developmental stage and limited exposure to advanced technologies [54], students intuitively leveraged the tools' distinct affordances in a sequential manner, using GPT for information seeking and DALL-E for visual inspiration (Finding 2). This structured approach highlights the students' ability to adapt to the unique strengths of these tools. GPT served as a textual guide, helping students address knowledge gaps and explore creative possibilities, while DALL-E supported visual ideation by translating abstract concepts into tangible representations. Importantly, this sequential usage was not explicitly taught; students independently aligned each tool with specific stages of their project workflows. This intuitive integration showcases the accessibility of generative AI tools for elementary learners, countering concerns about their potential opacity or complexity [38].

   Moreover, GenAI tools were proved effective in addressing the challenges elementary students face during the testing phase of PBL. Evidenced by significantly higher satisfaction levels reported by students using *AskArt* compared to the baseline group, GPT provided support for project information seeking and idea generation(Finding 1), bridging gaps in students' knowledge and sparking creativity. DALL-E complemented this by enabling students to visualize and refine abstract ideas (Finding 1), facilitating the translation of concepts into tangible representations. Together, these tools reduced the challenges often encountered during the ideation and implementation phases of group projects, which are frequently attributed to students' lack of knowledge and skills in teamwork and self-directed learning [8, 37]. Additionally, by offering immediate and individualized support, generative AIs acted as a valuable resource in classrooms where teachers were unable to

provide real-time assistance to all groups [19], enhancing students' ability to progress independently in their projects.

However, the challenge of formulating queries(Finding 3) highlights a critical gap in the accessibility of generative AI tools for younger users. While prompting challenges also occur in higher education settings, where users may struggle with optimizing specificity or creativity [39], the difficulties faced by elementary school students are fundamentally rooted in their developmental stages. Younger learners often lack the computational thinking skills and familiarity with AI systems required to align their goals with the affordances of these tools [29]. This gap limits their ability to navigate interfaces effectively and to grasp the iterative nature of prompting, leading to a greater reliance on external assistance, as observed in our findings (Finding 3), and a reliance on trial-and-error approaches.

As for groups using AI, the two collaboration strategies(Finding 4) observed in the student groups—assigning fixed members versus allowing everyone to interact with *AskArt*—also reflect different ways students navigated group dynamics. Fixed-member strategies facilitated structured collaboration by designating a specialist to mediate AI interactions. This approach introduced a built-in process for aligning the group on challenges and solutions, as the specialist gathered input from the team and reported AI-generated insights. However, it also centralized control over GenAI interactions, risking having personal bias for conversations with GenAI. On the other hand, allowing all students to interact freely with GenAI provided greater flexibility, enabling individuals to explore AI-generated suggestions independently. Yet, this approach did not necessarily enhance group engagement, particularly for students who were already less involved in collaboration. While the introduction of GenAI influenced collaboration dynamics, team engagement ratings (Figure 7 (b)) showed no significant difference between the experimental and baseline groups. This suggests that while GenAI offers opportunities to support collaboration—such as by facilitating consensus-building [31]—its use alone does not inherently lead to improved teamwork or engagement without deliberate scaffolding and coordination [20].

7.2 Design Considerations for GenAI-Powered Learning Tools Targeting K-6 Students

Based on the findings of RQ2, we propose three key considerations for designing GenAI-powered learning tools tailored to K-6 students.

7.2.1 Streamlining the input flow to GenAIs based on students' usage patterns. our findings revealed that students often used GPT and DALL-E in a sequential order(Finding 2). *AskArt* streamlines this flow with a "Select and Generate" option (Figure 5 A5), which enables students to select specific words or sentences from the chat with GPT and use them as the main subject or context words to generate images in DALL-E. It was proved to be helpful for help students iterate their queries to DALL-E(Finding 8). These findings are in line with previous research [23] showing that a structured input feature can assist users in constructing effective prompts to GenAIs. Therefore, we suggest that future GenAI-powered learning support tools should equip features that make it easy for elementary students to have subsequent queries to GenAIs in their learning activities. These tools can be inspired by previous work on chaining large language model (LLM) prompts, which has been shown to improve the quality of outcome and sense of human-AI collaboration [58]. For example, based on the user pattern of iteratively refining queries to DALL-E observed in our study, future tools could suggest prompts *e.g.,* about related objects, backgrounds, and goals, to improve the generated images in the last turn. Nevertheless, developers of these learning support tools should observe students' usage patterns in the target learning activities and iteratively improve their tools, which is a requirement in the user-centered design process.

7.2.2 Guiding Initial Use and Sustained Engagement with GenAIs. In Phases 3 and 4, we observed that providing context-specific example queries significantly improved students' understanding and proficiency in using GenAIs (Finding 5). For instance, the project-related self-introduction of *AskArt* and the example queries embedded in its interface (Figure 4 A2) were frequently adopted by students to initiate interactions. These use cases, preset by the research team to align with the specific themes and tasks of each course, enabled students to engage with GenAIs more confidently and effectively. Extending this feature, future tools could offer customizable interfaces for teachers to design example queries tailored to their classroom context. Templates based on project themes, student skill levels, and learning goals could be integrated with LLM capabilities to automatically generate relevant use cases, further supporting students in their initial interactions with GenAIs.

Beyond familiarity, keeping students engaged during conversations with GenAIs remains critical. The "Suggested Follow-Ups" feature in *AskArt* (Finding 7) demonstrated its potential to transform how students perceive and interact with GPT. We observed that elementary students tended to approach GPT with a one-turn query mindset, similar to how they might use a traditional search engine, missing the opportunity to leverage the conversational depth and iterative power unique to GenAIs. The suggested follow-ups helped introduce students to the concept of multi-turn dialogue, guiding them to build on previous responses and ask further questions. This shift not only facilitated deeper engagement, but also allowed students to explore more sophisticated, layered thinking through their interactions.

7.2.3 Scaffolding Query Articulation for Elementary Students. While the "Audio-to-Text Input" feature aimed to address challenges in typing and query formulation (Finding 3), it proved to be ineffective (Finding 6). Students' voice inputs were often disorganized or incomplete, resulting in unsatisfying responses from GPT. This underscores a more fundamental challenge: elementary students frequently struggle to clearly and correctly express their questions. Addressing this unique challenge requires solutions that go beyond simply mitigating typing difficulties. Future tools should focus on scaffolding students' ability to articulate their queries, starting with features that help them structure and clarify their thoughts. For example, tools could provide guided templates[12, 60] or visual prompts[45, 49] that break down complex queries into simpler components, helping students focus on their intent and the necessary details. Additionally, timely and user-adaptive support can proactively detect ambiguities or incomplete queries and engage users in interactive clarification and refinement[14, 62]. By tackling the root of these challenges, such tools could empower students to engage in more meaningful and productive interactions with GenAIs, ultimately enhancing their learning experience and fostering greater independence in their use of AI systems.

7.3 Teacher Roles in Leveraging GenAI for Effective PBL Classrooms

Based on the findings of RQ3, our findings indicate three considerations for teachers' roles in GenAI-enhanced PBL art courses.

7.3.1 Customizing GenAI Tools to Align with PBL Stages. Teachers' support for integrating GenAIs into PBL from the study (Finding 9) highlights the potential of these tools to enhance inquiry-driven learning. Beyond simply deploying GenAIs for information seeking and solution ideation, teachers could play an active role in customizing the settings of these tools to align with the specific stages of the PBL process. Customizations could include project-specific contexts, themes, or structured templates that guide students in crafting queries and reflecting on AI outputs. For instance, teachers could design prompts or configurations in GPT that encourage students to explore targeted concepts, themes, or methods related to their projects, enabling focused and guided exploration. Similarly, DALL-E settings could be adjusted to emphasize visual outputs that align with particular styles, techniques, or cultural elements central to the learning objectives. These adjustments ensure that

students' interactions with GenAIs remain both relevant and immersive, providing a structured yet flexible framework for exploration and ideation. This approach aligns with prior work that highlights how structured frameworks can help novices organize information, think critically, and improve creative ideas by embedding LLM capabilities into familiar templates[60].

7.3.2   Adapting GenAI Guidance to Student Abilities and Task Complexity. Teachers should tailor their guidance on GenAI usage to address the varying abilities of students while aligning it with the nature of the PBL task. Teachers noted that DALL-E provided a strong starting point for students with deficiencies in drawing or imagination, while GPT improved efficiency by organizing information and supporting task completion (Finding 10). These strengths make GenAIs particularly effective for abstract or complex tasks requiring extensive contextual understanding. For example, teachers could introduce pre-generated suggestions or scaffolded prompts to help less confident students engage with GenAIs, while advanced students could refine outputs or explore iterative design tasks requiring greater creative freedom. For simpler tasks—such as creating posters in Phase 4—conventional tools like search engines may suffice, allowing students to focus on creative ideation without the cognitive load of navigating GenAIs. In Phase 4, no significant difference was observed in perceived challenge levels between the experimental group (AskArt) and the control group, and the baseline group reported slightly higher creativity scores (Figure 7 (b)). However, for conceptually challenging projects, integrating GPT for synthesizing information and DALL-E for visual ideation can deepen inquiry and creativity. By aligning the use of GenAI with the complexity of tasks and student needs, teachers can foster inclusive collaboration while enhancing the PBL workflow.

7.3.3   Aligning GenAI Use with Task Nature and Student Abilities to Reduce Misuse. Teachers should ensure that students' use of GenAIs aligns with PBL learning objectives by addressing concerns about inappropriate use(Finding 11), such as over-reliance on AI outputs or superficial engagement . We recommend that GenAI integration consider the nature of the task and students' abilities. For tasks involving information seeking and ideation, GenAIs are highly effective, but students with stronger skills should be encouraged to go beyond replication and demonstrate originality and critical thinking. Conversely, students with limited skills might use AI-generated content as a foundation to develop their abilities. To address these concerns, teachers can require students to document their GenAI interactions, including prompts, outputs, and their decision-making processes. This aligns with Zheng et al. [64], who advocate for tracking task allocation and integration of AI suggestions to provide insight into student contributions. For tasks that could be completed easily with AI outputs, teachers might also require students to submit AI-generated materials alongside their final work to evaluate how they engaged with and adapted the outputs.

7.4   Limitations

Our work has several limitations. First, since the teaching assistants (TA) were not presented in the course sessions in Phase 1, and there were no recordings or audio captures, all qualitative feedback in Phase 1 was derived from the interviews with the teacher. This lack of comprehensive resources might have resulted in overlooking some potential design requirements for *AskArt*. Second, in Phase 4, no preliminary materials about GenAI were provided to the students, and they were given only 40 minutes to learn and use it. This short interaction time may have led to incomplete feedback from the students. Finally, from Phase 2 to Phase 4, the involvement of three TAs may create an environment that differs from a regular classroom setting in elementary schools. In the future, when students get familiar with the usage of GenAI, we could observe the students' behaviors without the presence of TAs in more art course sessions. Finally, our experimental design did not include a comprehensive baseline comparison. A more systematic approach, as conducted by prior

work[61], would involve equipping baseline groups with computer interfaces that allow the use of search engines for broader data collection.

8 Conclusion

In this study, we investigated the integration of GenAI in K-6 project-based art courses through a four-phase field study. Our research revealed the multifaceted impact of GenAIs, highlighting both their benefits and challenges. Students experienced enhanced information, inspiration, and guidance for their projects but struggled with query formulation to GenAI. The introduction of *AskArt*, a tailored interface combining DALL-E and ChatGPT, facilitated improved interaction patterns and increased engagement, yet indicated areas needing further refinement. Different collaboration strategies emerged among students, with increased engagement observed due to GenAIs. Teachers appreciated the enhanced student participation but voiced concerns about potential misuse and the need for more suitable interfaces for elementary school students. Based on the findings, our study further discussed key areas including information seeking, team collaboration, and the evolving roles of teachers in GenAI-enhanced classrooms. This research offers valuable insights into the effective use of GenAIs in K-6 project-based learning, presents *AskArt* as a practical solution for customizing GenAI interactions, and provides actionable recommendations for educators and researchers. These contributions aim to support the development and implementation of GenAI-powered learning tools in elementary education, advancing our understanding of their potential to transform student learning and engagement.

Acknowledgments

References

[1] Aoife A Ahern. 2010. A case study: Problem-based learning for civil engineering students in transportation courses. *European Journal of Engineering Education* 35, 1 (2010), 109–116.

[2] Roudha Saif Al-Khaldi, Doaa Elamin Emam, Yasser M Al Hamidi, Marwan Khraisheh, et al. 2023. Creating Creative Educational Opportunities among Engineering and Arts Students. In *2023 ASEE Annual Conference & Exposition*.

[3] Rakesh Belwal, Shweta Belwal, Azlinor Binti Sufian, and Amal Al Badi. 2020. Project-based learning (PBL): outcomes of students' engagement in an external consultancy project in Oman. *Education+ Training* 63, 3 (2020), 336–359.

[4] Karen E Bledsoe and Lawrence Flick. 2012. Concept development and meaningful learning among electrical engineering students engaged in a problem-based laboratory experience. *Journal of Science Education and Technology* 21 (2012), 226–245.

[5] Christy K Boscardin, Brian Gin, Polo Black Golde, and Karen E Hauer. 2024. ChatGPT and generative artificial intelligence for medical education: potential impact and opportunity. *Academic Medicine* 99, 1 (2024), 22–27.

[6] Thomas Breideband, Jeffrey Bush, Chelsea Chandler, Michael Chang, Rachel Dickler, Peter Foltz, Ananya Ganesh, Rachel Lieber, William R. Penuel, Jason G. Reitman, John Weatherley, and Sidney D'Mello. 2023. The Community Builder (CoBi): Helping Students to Develop Better Small Group Collaborative Learning Skills. In *Companion Publication of the 2023 Conference on Computer Supported Cooperative Work and Social Computing* (Minneapolis, MN, USA) *(CSCW '23 Companion)*. Association for Computing Machinery, New York, NY, USA, 376–380. https://doi.org/10.1145/3584931.3607498

[7] Bodong Chen, Xinran Zhu, et al. 2023. Integrating generative AI in knowledge building. *Computers and Education: Artificial Intelligence* 5 (2023), 100184.

[8] Juebei Chen, Anette Kolmos, and Xiangyun Du. 2021. Forms of implementation and challenges of PBL in engineering education: a review of literature. *European Journal of Engineering Education* 46, 1 (2021), 90–115.

[9] Liuqing Chen, Shuhong Xiao, Yunnong Chen, Yaxuan Song, Ruoyu Wu, and Lingyun Sun. 2024. ChatScratch: An AI-Augmented System Toward Autonomous Visual Programming Learning for Children Aged 6-12. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–19.

[10] Chin-Ling Chiang and Huei Lee. 2016. The effect of project-based learning on learning motivation and problem-solving ability of vocational high school students. *International Journal of Information and Education Technology* 6, 9 (2016), 709–712.

[11] Li-Yuan Chiou, Peng-Kai Hung, Rung-Huei Liang, and Chun-Teng Wang. 2023. Designing with AI: An exploration of co-ideation with image generators. In *Proceedings of the 2023 ACM designing interactive systems conference.* 1941–1954.

[12] Giulia Di Fede, Davide Rocchesso, Steven P Dow, and Salvatore Andolina. 2022. The idea machine: LLM-based expansion, rewriting, combination, and suggestion of ideas. In *Proceedings of the 14th Conference on Creativity and Cognition.* 623–627.

[13] Yogesh K Dwivedi, Laurie Hughes, Elvira Ismagilova, Gert Aarts, Crispin Coombs, Tom Crick, Yanqing Duan, Rohita Dwivedi, John Edwards, Aled Eirug, et al. 2021. Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management* 57 (2021), 101994.

[14] Pierre Erbacher, Ludovic Denoyer, and Laure Soulier. 2022. Interactive query clarification and refinement via user simulation. In *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval.* 2420–2425.

[15] Patrick Esser, Sumith Kulal, Andreas Blattmann, Rahim Entezari, Jonas Müller, Harry Saini, Yam Levi, Dominik Lorenz, Axel Sauer, Frederic Boesel, et al. 2024. Scaling rectified flow transformers for high-resolution image synthesis. In *Forty-first International Conference on Machine Learning.*

[16] Vivian L Gadsden. 2008. The arts and education: Knowledge generation, pedagogy, and the discourse of learning. *Review of research in education* 32, 1 (2008), 29–61.

[17] Susanne Garvis, Danielle Twigg, and Donna Pendergast. 2011. Breaking the negative cycle: The formation of self-efficacy beliefs in the arts. A focus on professional experience in pre-service teacher education. *Australasian Journal of Early Childhood* 36, 2 (2011), 36–41.

[18] Arnab Ghosh, Richard Zhang, Puneet K Dokania, Oliver Wang, Alexei A Efros, Philip HS Torr, and Eli Shechtman. 2019. Interactive sketch & fill: Multiclass sketch-to-image translation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision.* 1171–1180.

[19] Michail Giannakos, Roger Azevedo, Peter Brusilovsky, Mutlu Cukurova, Yannis Dimitriadis, Davinia Hernandez-Leo, Sanna Järvelä, Manolis Mavrikis, and Bart Rienties. 2024. The promise and challenges of generative AI in education. *Behaviour & Information Technology* (2024), 1–27.

[20] Yuanning Han, Ziyi Qiu, Jiale Cheng, and Ray Lc. 2024. When teams embrace AI: human collaboration strategies in generative prompting in a creative design task. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems.* 1–14.

[21] Ahmad Dasuki Mohd Hawari and Azlin Iryani Mohd Noor. 2019. Challenges In The Teaching Of Art Common Practices To Enhance Students' Soft Skills Via Project-Based Learning. In *8th UPI-UPSI International Conference 2018 (UPI-UPSI 2018).* Atlantis Press, 26–31.

[22] Ahmad Dasuki Mohd Hawari and Azlin Iryani Mohd Noor. 2020. Project based learning pedagogical design in STEAM art education. *Asian Journal of University Education* 16, 3 (2020), 102–111.

[23] Peiling Jiang, Jude Rayan, Steven P Dow, and Haijun Xia. 2023. Graphologue: Exploring large language model responses with interactive diagrams. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology.* 1–20.

[24] David R Johnson, Linda Renzulli, Jackson Bunch, and Maria Paino. 2013. Everyday observations: Developing a sociological perspective through a portfolio term project. *Teaching Sociology* 41, 3 (2013), 314–321.

[25] Kostas Karpouzis, Dimitris Pantazatos, Joanna Taouki, and Kalliopi Meli. 2024. Tailoring Education with GenAI: A New Horizon in Lesson Planning. *arXiv preprint arXiv:2403.12071* (2024).

[26] Majeed Kazemitabaar, Runlong Ye, Xiaoning Wang, Austin Zachary Henley, Paul Denny, Michelle Craig, and Tovi Grossman. 2024. Codeaid: Evaluating a classroom deployment of an llm-based programming assistant that balances student and educator needs. In *Proceedings of the CHI Conference on Human Factors in Computing Systems.* 1–20.

[27] Jingoog Kim and Mary Lou Maher. 2023. The effect of AI-based inspiration on human design ideation. *International Journal of Design Creativity and Innovation* 11, 2 (2023), 81–98.

[28] Dimitra Kokotsaki, Victoria Menzies, and Andy Wiggins. 2016. Project-based learning: A review of the literature. *Improving schools* 19, 3 (2016), 267–277.

[29] Irene Lee, Safinah Ali, Helen Zhang, Daniella DiPaola, and Cynthia Breazeal. 2021. Developing middle school students' AI literacy. In *Proceedings of the 52nd ACM technical symposium on computer science education.* 191–197.

[30] Pei-Ying Lin, Kristina Andersen, Ralf Schmidt, Sanne Schoenmakers, Hèrm Hofmeyer, Pieter Pauwels, and Wijnand IJsselsteijn. 2024. Text-to-Image AI as a Catalyst for Semantic Convergence in Creative Collaborations. In *Proceedings of the 2024 ACM Designing Interactive Systems Conference.* 2753–2767.

[31] Pei-Ying Lin, Kristina Andersen, Ralf Schmidt, Sanne Schoenmakers, Hèrm Hofmeyer, Pieter Pauwels, and Wijnand IJsselsteijn. 2024. Text-to-Image AI as a Catalyst for Semantic Convergence in Creative Collaborations. In *Proceedings of the 2024 ACM Designing Interactive Systems Conference*. 2753–2767.

[32] Shih-Yun Lu, Chih-Cheng Lo, and Jia-Yu Syu. 2022. Project-based learning oriented STEAM: The case of micro–bit paper-cutting lamp. *International Journal of Technology and Design Education* 32, 5 (2022), 2553–2575.

[33] Galyna Lutsenko. 2018. Case study of a problem-based learning course of project management for senior engineering students. *European Journal of Engineering Education* 43, 6 (2018), 895–910.

[34] Milan Maros, Marcela Korenkova, Milan Fila, Michal Levicky, and Maria Schoberova. 2023. Project-based learning and its effectiveness: evidence from Slovakia. *Interactive Learning Environments* 31, 7 (2023), 4147–4155.

[35] Siva Mathiyazhagan and Karolina La Fors. 2023. Children's right to participation in AI: Exploring transnational co-creative approaches to foster child-inclusive AI policy and practice. *Information Polity* 28, 1 (2023), 141–153.

[36] Michael B McCullough, Michael D Martin, and Mollika A Sajady. 2018. Implementing green walls in schools. *Frontiers in psychology* 9 (2018), 357834.

[37] Guien Miao, Iresha Ranaraja, Sarah Grundy, Nick Brown, Marina Belkina, and Tom Goldfinch. 2024. Project-based learning in Australian & New Zealand universities: current practice and challenges. *Australasian Journal of Engineering Education* (2024), 1–13.

[38] Michele Newman, Kaiwen Sun, Ilena B Dalla Gasperina, Grace Y. Shin, Matthew Kyle Pedraja, Ritesh Kanchi, Maia B. Song, Rannie Li, Jin Ha Lee, and Jason Yip. 2024. "I want it to talk like Darth Vader": Helping Children Construct Creative Self-Efficacy with Generative AI. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 117, 18 pages. https://doi.org/10.1145/3613904.3642492

[39] Srishti Palani, Zijian Ding, Austin Nguyen, Andrew Chuang, Stephen MacNeil, and Steven P Dow. 2021. CoNotate: Suggesting queries based on notes promotes knowledge discovery. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–14.

[40] Zachary A Pardos, Matthew Tang, Ioannis Anastasopoulos, Shreya K Sheel, and Ethan Zhang. 2023. Oatutor: An open-source adaptive tutoring system and curated content library for learning sciences research. In *Proceedings of the 2023 chi conference on human factors in computing systems*. 1–17.

[41] Victoria Pavlou and Aurelio Castro-Varela. 2024. E-Learning Canvases: Navigating the Confluence of Online Arts Education and Sustainable Pedagogies in Teacher Education. *Sustainability* 16, 5 (2024), 1741.

[42] Xiaohan Peng, Janin Koch, and Wendy E Mackay. 2024. Designprompt: Using multimodal interaction for design exploration with generative ai. In *Proceedings of the 2024 ACM Designing Interactive Systems Conference*. 804–818.

[43] Loginova Polina, Klakova Vlada, Erik Bohemia, Tatiana Semichevskaya, Anna Solovyova, et al. 2022. Student experience of online international design studio participation. In *DS 117: Proceedings of the 24th International Conference on Engineering and Product Design Education (E&PDE 2022), London South Bank University in London, UK. 8th-9th September 2022*.

[44] Aditya Ramesh, Mikhail Pavlov, Gabriel Goh, Scott Gray, Chelsea Voss, Alec Radford, Mark Chen, and Ilya Sutskever. 2021. Zero-shot text-to-image generation. In *International conference on machine learning*. Pmlr, 8821–8831.

[45] Aneeshan Sain, Ayan Kumar Bhunia, Yongxin Yang, Tao Xiang, and Yi-Zhe Song. 2021. Stylemeup: Towards style-agnostic sketch-based image retrieval. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 8504–8513.

[46] Mariela Samaniego, Nancy Usca, José Salguero, and William Quevedo. 2024. Creative Thinking in Art and Design Education: A Systematic Review. *Education Sciences* 14, 2 (2024), 192.

[47] Orit Shaer, Angelora Cooper, Osnat Mokryn, Andrew L Kun, and Hagit Ben Shoshan. 2024. AI-Augmented Brainwriting: Investigating the use of LLMs in group ideation. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–17.

[48] Mike Sharples. 2023. Towards social generative AI for education: theory, practices and ethics. *Learning: Research and Practice* 9, 2 (2023), 159–167.

[49] Kihoon Son, DaEun Choi, Tae Soo Kim, Young-Ho Kim, and Juho Kim. 2024. Genquery: Supporting expressive visual search with generative models. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. 1–19.

[50] Yuqian Sun, Ying Xu, Chenhang Cheng, Yihua Li, Chang Hee Lee, and Ali Asadipour. 2023. Explore the Future Earth with Wander 2.0: AI Chatbot Driven By Knowledge-base Story Generation and Text-to-image Model. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI EA '23)*. Association for Computing Machinery, New York, NY, USA, Article 450, 5 pages. https://doi.org/10.1145/3544549.3583931

[51] Jianqing Tang and Guoxia Zou. 2023. Taking the Software Technology Specialty as an Example to Construct the Professional Knowledge List with Project-Based Learning Method. *Open Journal of Social Sciences* 11, 12 (2023), 27–34.

[52] Jakob Tholander and Martin Jonsson. 2023. Design ideation with AI-sketching, thinking and talking with generative machine learning models. In *Proceedings of the 2023 ACM designing interactive systems conference*. 1930–1940.

[53] U Usmeldi. 2019. The Effect of Project-based Learning and Creativity on the Students' Competence at Vocational High Schools, 299, 14–17.

[54] Maarten Van Mechelen, Rachel Charlotte Smith, Marie-Monique Schaper, Mariana Tamashiro, Karl-Emil Bilstrup, Mille Lunding, Marianne Graves Petersen, and Ole Sejer Iversen. 2023. Emerging technologies in K–12 education: A future HCI research agenda. *ACM Transactions on Computer-Human Interaction* 30, 3 (2023), 1–40.

[55] Maarten Van Mechelen, Rachel Charlotte Smith, Marie-Monique Schaper, Mariana Tamashiro, Karl-Emil Bilstrup, Mille Lunding, Marianne Graves Petersen, and Ole Sejer Iversen. 2023. Emerging Technologies in K–12 Education: A Future HCI Research Agenda. *ACM Trans. Comput.-Hum. Interact.* 30, 3, Article 47 (jun 2023), 40 pages. https://doi.org/10.1145/3569897

[56] Mathias Peter Verheijden and Mathias Funk. 2023. Collaborative Diffusion: Boosting Designerly Co-Creation with Generative AI. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI EA '23)*. Association for Computing Machinery, New York, NY, USA, Article 73, 8 pages. https://doi.org/10.1145/3544549.3585680

[57] Brett R Wildermoth, David D Rowlands, et al. 2012. Project based learning in embedded systems: A case study. In *Proceedings of the 2012 AAEE Conference. Melbourne, Victoria, Australia*. Citeseer.

[58] Tongshuang Wu, Michael Terry, and Carrie Jun Cai. 2022. Ai chains: Transparent and controllable human-ai interaction by chaining large language model prompts. In *Proceedings of the 2022 CHI conference on human factors in computing systems*. 1–22.

[59] Ceisy N Wuntu, Yolanda Singal, and Olga A Rorintulus. 2022. The Implementation of Project Based Learning (PBL) in Improving Students' Speaking Skill at SMA Yadika Kopandakan II. *International Journal of English and Applied Linguistics (IJEAL)* 2, 3 (2022), 387–398.

[60] Xiaotong Xu, Jiayu Yin, Catherine Gu, Jenny Mar, Sydney Zhang, Jane L E, and Steven P Dow. 2024. Jamplate: Exploring LLM-Enhanced Templates for Idea Reflection. In *Proceedings of the 29th International Conference on Intelligent User Interfaces*. 907–921.

[61] Yilin Ye, Qian Zhu, Shishi Xiao, Kang Zhang, and Wei Zeng. 2024. The contemporary art of image search: Iterative user intent expansion via vision-language model. *Proceedings of the ACM on Human-Computer Interaction* 8, CSCW1 (2024), 1–31.

[62] Michael JQ Zhang and Eunsol Choi. 2023. Clarify when necessary: Resolving ambiguity through interaction with lms. *arXiv preprint arXiv:2311.09469* (2023).

[63] Xiaohong Zhang and Kenichi Kubota. 2023. Cultivating Students' Creative Thinking Using Visual Narrative in an Agile Blended Learning Environment. In *International Conference on Blended Learning*. Springer, 163–173.

[64] Chengbo Zheng, Kangyu Yuan, Bingcan Guo, Reza Hadi Mogavi, Zhenhui Peng, Shuai Ma, and Xiaojuan Ma. 2024. Charting the Future of AI in Project-Based Learning: A Co-Design Exploration with Students. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–19.

[65] Xiaofei Zhou, Jessica Van Brummelen, and Phoebe Lin. 2020. Designing AI learning experiences for K-12: emerging works, future opportunities and a design framework. *arXiv preprint arXiv:2009.10228* (2020).

[66] Qingxiaoyang Zhu and Hao-Chuan Wang. 2023. Leveraging Large Language Model as Support for Human Problem Solving: An Exploration of Its Appropriation and Impact. In *Companion Publication of the 2023 Conference on Computer Supported Cooperative Work and Social Computing* (Minneapolis, MN, USA) *(CSCW '23 Companion)*. Association for Computing Machinery, New York, NY, USA, 333–337. https://doi.org/10.1145/3584931.3606965

[67] Yan Zhuge and Julie E Mills. 2009. Teaching finite element modelling at the undergraduate level: a PBL approach. In *Proceedings of the 20th Annual Conference of the Australasian Association for Engineering Education (AAEE 2009)*.