

The Impact of Contemporary Art Appreciation Teaching on the Cultivation of Critical Thinking in Chinese Middle School Students

YuLei Wang¹, Mo Yuan², Yang Wang², Guanbo Wang²,
Hongbo Wang², Zhongqi Yao², Hoi Yeh Lee^{1*}

¹Universiti Pendidikan Sultan Idris, Tanjung Malim, Perak, Malaysia, 35900, ²School of Computer and Information Engineering, Heihe University, Heilongjiang, China, 164300
Email: wangyulei007@126.com

DOI Link: <http://dx.doi.org/10.6007/IJARBS/v16-i9/28975>

Published Date: 29 September 2026

Abstract

In the context of competency-oriented basic education, art education is gradually shifting from the transmission of artistic knowledge and skills toward the development of students' comprehensive abilities. The *Compulsory Education Art Curriculum Standards (2022 Edition)* emphasizes the educational value of art and the cultivation of students' aesthetic perception, artistic expression, creative practice, and cultural understanding. At the same time, the rapid development of visual culture and digital media has exposed middle school students to increasingly diverse images and information, making their abilities to analyse information, evaluate evidence, form judgements, and reflect on their own thinking increasingly important. Contemporary art, characterized by openness, ambiguity, diverse media, and connections with social issues, provides a particularly suitable context for cultivating these abilities. Unlike appreciation approaches that emphasize predetermined interpretations, contemporary art can create cognitive conflicts and encourage students to ask questions, seek evidence, compare different perspectives, construct arguments, and reconsider their initial judgements. However, the extent to which structured contemporary art appreciation teaching can systematically promote critical thinking among Chinese middle school students remains insufficiently explored. Therefore, this study investigates the impact of contemporary art appreciation teaching on the critical thinking of Grade 8 Chinese middle school students. Two parallel classes, with 30 students in each, are included in a quasi-experimental design. The experimental group receives structured contemporary art appreciation teaching, while the control group receives conventional art appreciation instruction. Students' critical thinking is examined across five dimensions: question awareness, analytical ability, reasoning ability, evaluative ability, and reflective ability. The teaching process follows the sequence of observation → questioning → evidence → argumentation → discussion and evaluation → reflection, with four themes focusing on what can be considered art, how art expresses social issues, whether artworks can have different interpretations, and how contemporary artworks

can be evaluated. In addition, Qwen3-8B is introduced as an auxiliary evaluation tool to analyse quantitative scores and classroom discourse. The study aims to examine both the effectiveness and underlying mechanisms of contemporary art appreciation teaching, develop a practical teaching model, and explore the potential and limitations of large language models in educational assessment. (Ministry of Education of the People's Republic of China, 2022)

Keywords: Contemporary Art, Artwork Appreciation, Middle School Art, Critical Thinking, Core Competencies, Large Language Model-Assisted Evaluation

Introduction

Research Background

Against the backdrop of the rapid development of digital media and the widespread prevalence of visual culture, the ways in which adolescents acquire information are undergoing significant shifts. Images, short-form videos, cyber-visual symbols, and AI-generated content have become integral components of students' daily information environment. Compared to traditional text-based information, visual information is typically characterized by rapid dissemination, complex forms, and open-ended meanings. Students are required not merely to "see" visual data, but also to evaluate information sources, decode visual symbols, identify discrepancies among competing interpretations, and form sound judgments grounded in relevant evidence. Consequently, how to cultivate students' capacities to analyze information, evaluate evidence, construct arguments, and engage in reflective judgment during basic education has emerged as a pressing educational concern.

Critical thinking does not equate to simply refuting others' perspectives; rather, it underscores an individual's ability to analyze an issue, substantiate judgments with evidence, and revise preconceived notions in light of new information. Existing literature suggests that critical thinking encompasses both cognitive skills—such as analysis, inference, and evaluation—and intellectual dispositions, including open-mindedness, prudence, and reflexivity. Its cultivation demands sustained instructional activities situated within concrete disciplinary contexts. A 2023 systematic review further noted that while definitions of critical thinking vary, scholarly consensus treats it as a higher-order cognitive capability developable through instructional intervention, while also pointing out substantial room for inquiry regarding effective pedagogical strategies and assessment models. (Paul & Elder, 2013; Zhong, 2015)

Within this context, arts education holds distinct pedagogical value. Artworks are not merely aesthetic objects; they can also function as cognitive material for observation, interpretation, comparison, and evaluation. Particularly in visual arts learning, students grapple with intricate relationships among form, medium, symbolism, cultural contexts, and artistic intent—a process deeply intertwined with critical thinking operations such as analysis, inference, evaluation, and reflection. Recent research has also begun to re-examine visual arts education through the lenses of cognitive development and learning modes. For instance, studies have revealed that visual arts learning fosters habits of mind such as sustained observation, analysis, decision-making, and reflection; Visual Thinking Strategies (VTS) have likewise been utilized to train learners in observation, interpretation, and evidence-based reasoning. (Efland, 2004; Wang, 2017)

Compared to traditional canonical art, contemporary art exhibits far more pronounced openness, polysemy, and social-issue orientation. Works frequently engage real-world issues such as the environment, consumerism, technology, identity, and urban life, the meanings of which resist reduction to monolithic art-historical conclusions. Such interpretive indeterminacy creates rich space for questioning, evidence analysis, perspective comparison, and argumentation in the classroom. A 2023 study on aesthetic inquiry indicated that in practical visual arts classrooms, teaching still tends to gravitate toward formal analysis and technical skill acquisition, leaving students' exploration and construction of artistic meaning relatively underdeveloped. This highlights an urgent need to examine how art appreciation can be further transformed into an inquiry-based learning process. (Qian, 2013; Zhu, 2019; Li, 2021)

Moreover, recent scholarship in arts education has increasingly adopted quasi-experimental methods to examine student-centered instructional models. For example, a study on choice-based art education in secondary schools employed an experimental-control group design to examine the impacts of art pedagogy on outcome variables such as innovative thinking, signaling an empirical turn in arts education research away from purely descriptive accounts of teaching experience. (Wang, 2017)

Nevertheless, several gaps in the literature warrant further discussion. First, while research examining the nexus between arts education and higher-order thinking is expanding, empirical studies specifically focused on the intersection of Chinese junior secondary students, contemporary art appreciation, and critical thinking remain scarce. Second, existing inquiries have predominantly concentrated on the impact of arts education on creativity, learning interest, or aesthetic capability; empirical scrutiny into specific critical thinking dimensions—such as problem awareness, analysis, inference, evaluation, and reflection—requires greater granularity. Third, validating how the open nature of contemporary art can be operationalized into reproducible, observable, and assessable classroom instructional procedures demands further classroom-based experimental validation.

Accordingly, focusing on eighth-grade students in China, this study employs a quasi-experimental design to compare the effects of structured contemporary art appreciation instruction against conventional art appreciation pedagogy. It evaluates the development of students' critical thinking across five specific dimensions: problem awareness, analytical ability, inferential reasoning, evaluative competence, and reflective capacity. Furthermore, this study explores the feasibility of employing large language models (LLMs) to assist in educational assessment. In doing so, this study bridges the fields of arts education, critical thinking development, and AI-enabled educational assessment.

Research Significance

Theoretically, this paper combines critical thinking education with contemporary art appreciation at the middle school level, further illustrating that art appreciation has not only aesthetic-educational value but also cognitive-educational value. Practically, the research can help art teachers transform the teacher-centred classroom structure and shift students from 'answer receivers' to 'meaning constructors.' Existing research on visual art appreciation at the middle school level has shown that integrating online resources with critical thinking instruction — through pre-visit, on-site, and post-visit activities involving observation,

discussion, and reflection — offers valuable practical insights for the present study. (Efland, 2004; Yin, 2015)

Research Questions

This paper primarily addresses three questions:

1. Through what mechanisms can contemporary art appreciation teaching promote critical thinking in middle school students?
2. Compared with conventional art appreciation instruction, can structured contemporary art appreciation teaching improve students' skills in analysis, reasoning, evaluation, and reflection?
3. How can teachers reduce the difficulty of classroom implementation through question design, evidence analysis, collaborative discussion, and diversified assessment?

Related Work

Contemporary Art Appreciation Teaching

As used in this paper, contemporary art appreciation teaching refers to a pedagogical approach in which teachers select contemporary art cases that are timely, open-ended, and problem-oriented, guiding students through activities such as observing the formal aspects of works, analysing visual symbols, exploring creative contexts, raising questions, gathering evidence, comparing perspectives, and forming evaluative judgements to achieve understanding and interpretation of artworks. The emphasis is not on having students memorise the names of artists and movements, but on leading them through the thinking process of 'why do I understand it this way.'

Critical Thinking

Critical thinking can be understood as the capacity of an individual, when confronted with a problem or information, to actively raise questions, analyse the relationships between evidence and arguments, identify different viewpoints, engage in sound reasoning, form judgements according to standards, and simultaneously reflect on and revise those judgements. For middle school students, critical thinking does not mean rejecting everything; rather, it emphasises 'doubt grounded in reason,' 'judgement grounded in evidence,' and the ability to 'explain one's own reasoning.' (Paul & Elder, 2013; Zhong, 2015)

Theoretical Foundations

First, constructivist learning theory holds that learning is not a matter of teachers directly transmitting fixed answers to students, but of learners actively constructing meaning on the basis of prior experience through context, interaction, and problem-solving. The multiple meanings inherent in contemporary artworks can generate cognitive conflict, impelling students to reorganise their existing understanding. (Dewey, 1990)

Second, inquiry-based learning emphasises starting from authentic problems and having students experience the full cycle of raising questions, gathering information, analysing evidence, interpreting results, and reflecting on the process. Art appreciation is inherently inquiry-oriented and can therefore be naturally combined with critical thinking training. (Dewey, 1990)

Third, aesthetic education theory emphasises that students form aesthetic judgements in the process of perceiving, experiencing, understanding, and evaluating art. Related research suggests that high-quality classroom questioning can promote middle school students' deeper understanding of artworks and enhance their aesthetic capacities. (Yin, 2015; Zhao, 2020)

Overemphasis on Knowledge Recall in Teaching Objectives

Some classrooms still treat knowledge about artists' names, titles, dates of creation, and artistic movements as the primary objective, requiring students merely to memorise 'standard answers' while rarely asking them to explain why they formed a particular view. Knowledge itself is undeniably important, but without analytical and evaluative components, students find it difficult to convert knowledge into cognitive capacity. (Qian, 2013; Zhu, 2019)

Works Chosen Are Distant from Students' Lives

Traditional appreciation classes tend to focus on canonical works whose historical evaluations are already well established, and teachers often come prepared with ready-made conclusions. Contemporary art, by contrast, is more closely tied to social reality; yet some teachers worry that such works are 'too difficult to understand' and therefore rarely incorporate them into middle school lessons. This reduces students' opportunities to engage with open-ended visual problems. (Qian, 2013; Zhu, 2019)

Question Design Lacks Depth and Progression

Low-level questions such as 'What do you see in the picture?' or 'What colours were used?' help with observation, but if the lesson remains at this level, it is difficult to develop reasoning and evaluative abilities. Questions that genuinely promote critical thinking should press further: 'Why do you make that judgement?' 'Which part of the work supports your view?' 'If we tried a different interpretation, what evidence would change?' (Zhao, 2020; Wang, 2017)

Classroom Assessment Tends to Seek a Single Correct Answer (Zhao, 2020; Li, 2021)

Contemporary artworks often permit multiple interpretations. If teachers use 'how closely does the response match the textbook' as an evaluation standard, students may gradually develop a habit of relying on authoritative answers. It is therefore necessary to shift from outcome-based assessment to process-based assessment, giving weight to the questions students raise, the evidence they use, the soundness of their reasoning, and their ability to respond to differing viewpoints.

Methodology

Openness Fosters Question Awareness

Contemporary art frequently breaks away from traditional media and modes of representation, so students cannot immediately apply existing answers when they encounter unfamiliar works. For example, a work made from everyday objects might prompt the question 'Why can an ordinary object be art?' Teachers can hold that question open rather than rushing to provide a conclusion, setting students in a state of active inquiry.

Ambiguity Fosters Perspective Comparison

The same work may yield different interpretations. Teachers can organise students into groups with divergent views, requiring each group to articulate its position, evidence, and

reasoning process. Through comparison, students discover that a viewpoint is not the same as a fact, and that personal impressions must be supported by evidence drawn from the work and contextual information.

Social Issues Foster Argumentation Skills

Contemporary art frequently addresses real-world issues such as the environment, urban life, technology, and consumption. Teachers can connect the work to students' lives, guiding them to answer questions such as 'What problem does this work raise?', 'How does the artist convey it?', and 'Is this form of expression effective?' In this way, students are required to move from description to analysis, and from analysis to evaluation.

Uncertainty Fosters Reflective Ability

There is not necessarily a single correct answer in art appreciation. After discussion, students can revisit their initial judgements, explaining 'What was my original view?' and 'What evidence changed my mind?' This cognitive revision is itself an important manifestation of critical thinking.

On this basis, a relatively complete causal chain can be constructed: the open context of contemporary art → cognitive conflict → question-raising → gathering and evaluating evidence → comparing perspectives → forming an argument → responding to challenges → reflective revision. This process is highly consistent with the core activities of critical thinking.

Research Design

Research Subjects

This study focused on two parallel Grade 8 classes at a given school: the experimental group (Class 1, Grade 8) and the control group (Class 2, Grade 8), each comprising 30 students, for a total of 60 participants. Prior to the study, pre-tests were used to compare the baseline critical thinking levels of the two classes. The experimental group's pre-test mean was 9.13 and the control group's was 9.07 (out of 25 points), with no significant between-group difference, confirming a balanced baseline suitable for between-group comparison.

Research Methods

This study adopted a quasi-experimental design combined with a mixed-methods approach incorporating questionnaires, artwork analysis, classroom observation, and interviews. The experimental class received instruction integrating 'contemporary art appreciation + critical thinking,' while the control class followed conventional appreciation instruction. Both groups completed pre- and post-tests, and the instructional effect was compared through before-and-after changes. In the data-analysis stage, the Qwen3-8B large language model was introduced to provide auxiliary validation of the student scoring data, exploring the validity of applying large language models in educational assessment.

Teaching Content

The instructional unit can be organised around four themes: Unit 1, 'What Can Be Art?', focusing on artistic concepts and media; Unit 2, 'How Does Art Express Social Issues?', focusing on the relationship between art and social reality; Unit 3, 'Can One Work Have Different Answers?', focusing on multiple interpretations; and Unit 4, 'How Do I Evaluate a

Contemporary Artwork?', providing integrated training in evidence, argumentation, and reflection.

Experimental Group Teaching Procedure

Phase 1 — Observation: Students independently record the visual facts present in the work without rushing to interpret them. Phase 2 — Questioning: Students raise questions about aspects they genuinely find puzzling. Phase 3 — Evidence: Students search for materials in the work's formal qualities, title, information about the artist, and social context to support their views. Phase 4 — Argumentation: Students express their position using the 'claim–evidence–reasoning' structure. Phase 5 — Debate and Peer Review: Students respond to alternative viewpoints. Phase 6 — Reflection: Students re-evaluate their initial positions. This model transforms the appreciation process from 'view a work → listen to an explanation' to 'view a work → raise questions → gather evidence → form a judgement.'

Evaluation Indicators

This study divides critical thinking into five dimensions: question awareness, analytical ability, reasoning ability, evaluative ability, and reflective ability. Each dimension is assessed on a 1–5 Likert scale, for a maximum total of 25 points. For example, a student who actively raises questions related to the work and of genuine inquiry value would receive a high rating on the question-awareness dimension; a student who cites specific details from the work to support a judgement would score highly on the evidence-analysis dimension; and a student who compares two or more interpretations and explains the reasons would demonstrate further development of evaluative and reasoning abilities.

Data Analysis

Quantitative data can be analysed using means, standard deviations, and pre-post differences; where conditions allow, analysis of covariance can further control for pre-test differences. Qualitative data focus on students' classroom utterances, learning worksheets, and written artwork evaluations, tracing the shift in thinking from 'I like it / I don't like it' to 'I believe ... because the work demonstrates ...'. It should be emphasised that this paper constitutes a research design and theoretical analysis; no specific experimental data have been fabricated. When submitting a formal paper, actual experimental data should replace the illustrative analysis.

Large Language Model-Assisted Evaluation

Building on human scoring, this study introduced Qwen3-8B to conduct auxiliary analysis of Dataset 1 (pre- and post-test scores for 60 students) and Dataset 2 (18 qualitative classroom observation records). The model independently scored each set of ratings according to the five-dimensional scale and made thinking-level judgements on classroom utterances (descriptive layer → analytical layer → reasoning layer → evaluative layer → reflective layer). Through a human-machine comparison, the practical utility and limitations of large language models in educational assessment contexts were assessed.

To illustrate with the case 'What Is Art: Can an Ordinary Object Become Art?': The teacher first presents an artwork made from a commonplace object removed from its everyday context of use and asks students to write down their immediate reactions within one minute. Rather than immediately revealing the work's background, the teacher poses three

progressively more demanding questions: 'What do you see?', 'What makes you think it might be art?', and 'If the artist had not told you the work's meaning, could you propose your own interpretation?'

The second phase involves evidence analysis. Working in groups, students record their observations from the perspectives of materials, scale, arrangement, display space, and title. The teacher asks students to distinguish between 'observed facts' and 'personal interpretations': for instance, 'the object is placed at the centre of the gallery' is a fact, whereas 'the artist is satirising consumer society' is an interpretation — the two must not be conflated.

In the third phase, students argue for their positions. They may choose from three stances — 'it is art,' 'it is not art,' or 'I cannot yet decide' — but must provide at least two pieces of supporting evidence. Other students may challenge the evidence. The teacher does not evaluate which position is absolutely correct; instead, the teacher assesses whether the argument is clear, whether the evidence is relevant, and whether the reasoning is sound.

In the fourth phase, information about the artist and the work's creation is provided, and students revise their judgements. They compare their two positions — before and after the new information — and fill in 'What changed in my view?' and 'What evidence influenced my judgement?' This process helps students recognise that a judgement can be adjusted as new evidence emerges, and that changing one's mind is not a failure.

The core purpose of this case is not to have students memorise an art-historical conclusion about a particular work, but to lead them through a complete cognitive process. Such a lesson embeds critical thinking within art learning itself, rather than imposing an additional layer of training unrelated to art.

Results and Discussion

Comparison of Pre- and Post-Test Average Scores across Five Dimensions of Critical Thinking

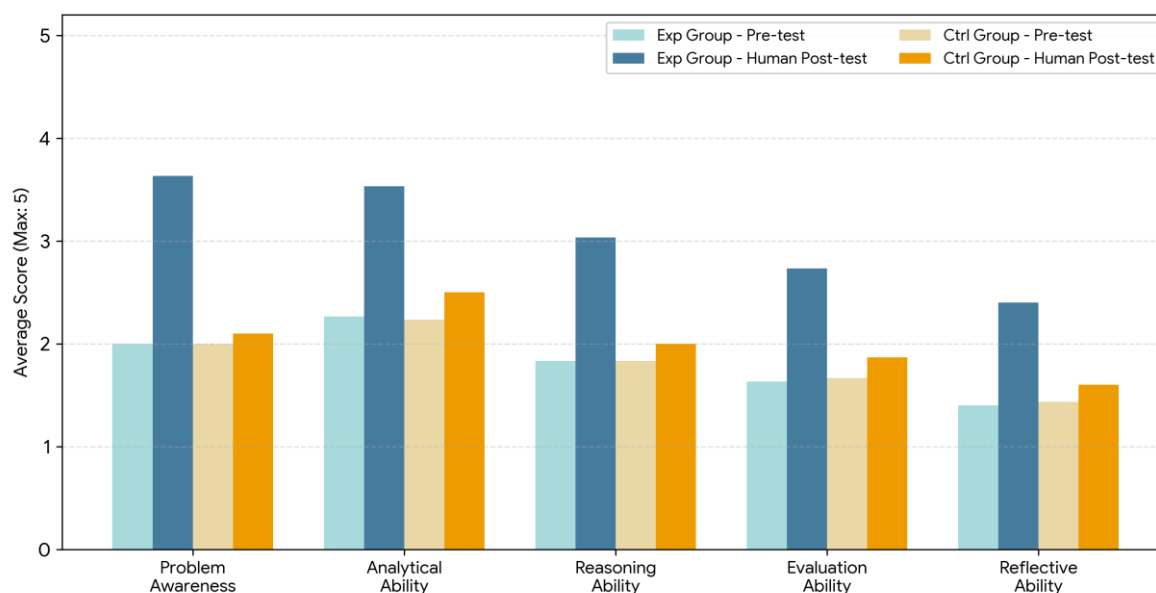


Chart 1

Quantitative Results

After the eight-lesson experimental intervention, changes in critical thinking scores for the experimental and control groups are summarised in the table below:

Summary of Pre- and Post-Test Critical Thinking Total Scores (Full Score: 25)

Group	Pre-test Mean	Pre-test SD	Post-test Mean (Human)	Post-test Mean (Qwen)	Mean Gain (Human)	Mean Gain (Qwen)	MAE	r
Experimental (n=30)	9.13	2.73	15.33	14.10	6.20	4.97	1.70	0.876
Control (n=30)	9.07	2.71	10.07	10.20	0.90	1.03	0.53	0.959

The experimental results show that the experimental class demonstrated significant improvements over the pre-test in critical thinking total scores and in the dimensions of analysis, reasoning, evaluation, and reflection, and performed significantly better than the control class — which received conventional lecture-based appreciation instruction — with a mean gain of 6.20 compared with 0.90. Qwen3-8B model scores were highly correlated with human scores (experimental group $r = 0.876$; control group $r = 0.959$), confirming the reliability of the human ratings and demonstrating that the large language model possesses basic validity as an auxiliary educational assessment tool. The primary reason is not that 'contemporary art naturally enhances thinking,' but that teachers used the openness of contemporary art to design more learning activities requiring students to observe, question, provide evidence, and construct arguments.

First, contemporary art can expand the problem space. Traditional works can usually be pre-explained by teachers, whereas the media, concepts, and social themes of contemporary art can create more pronounced cognitive conflict, generating authentic questions in students. Data show that the experimental group's mean gain on the 'question awareness' dimension (+1.37) was markedly higher than that of the control group (+0.13).

Second, evidence awareness is one of the most notable changes. Students initially tended to express judgements using phrases such as 'I feel' or 'I like,' but after training, gradually learned to cite specific visual details, titles, materials, and background information to support their views. Qualitative data show that student E001 progressed from 'I don't know if this counts as art' (descriptive layer) in Lesson 1 to 'The evidence is the labour embodied in 100 million porcelain seeds; my judgement is that it is questioning mass production' (reasoning + evaluative layer) by Lesson 7, reflecting a significant leap in thinking level.

Third, diversified evaluation increased cognitive openness. After discovering that the same work can yield different interpretations, students gradually recognised that 'different' does not mean 'all equally valid,' and that the quality of evidence and the soundness of reasoning require further judgement.

Fourth, the reflection phase helped students develop more stable thinking habits. Students became aware that their views might be shaped by prior impressions, and were willing to adjust their judgements in light of new evidence. Student E002 wrote in Lesson 8: 'I realise I

learned one thing in this unit: a good argument needs multiple pieces of evidence that support each other, not just one,' demonstrating typical metacognitive ability.

Qualitative Results and Qwen Level Judgements

Qualitative observation records from Dataset 2 reveal a systematic developmental trend in the experimental group's thinking levels. Taking student E001 as an example, their thinking progressed along the complete trajectory: descriptive layer (Lessons 1–2) → analytical layer (Lessons 3–4) → reasoning layer (Lesson 5) → evaluative layer (Lesson 6) → reasoning + evaluative layer (Lesson 7) → reflective layer (Lesson 8). Qwen3-8B's agreement rate on thinking-level judgements for two representative experimental-group cases (E001, Lesson 7; E002, Lesson 7) reached 100%, with an argumentation-completeness MAE of 1.0 (out of 5), indicating a sound capacity to identify high-level critical thinking performance.

Control-group students (exemplified by C001) largely remained at the descriptive layer throughout the entire experimental period. Their classroom utterances consisted mainly of restating the teacher's conclusions (e.g., 'The teacher said this is anti-war'), with little independent questioning or citation of evidence; their mean gain was 0.90, significantly lower than that of the experimental group.

It should be noted that contemporary art appreciation also has certain limitations. If the works are overly complex, if the teacher's background knowledge is insufficient, or if discussion lacks structure, the classroom may descend into 'anything can be an interpretation.' Openness must therefore coexist with evidence-based constraints. Teachers must neither restrict interpretation to a single standard answer nor treat all subjective impressions as equally valid. A classroom rule of 'open viewpoints + evidence-based constraints' needs to be established.

Transforming the Teacher's Role

Teachers should shift from being 'announcers of the work's meaning' to 'organisers of problem contexts' and 'guides of thinking processes.' When confronted with different student answers, teachers should prioritise asking for evidence and reasons rather than immediately judging correctness.

Establishing a Progressive Question Chain

Questions can be arranged progressively along the sequence: observe → understand → analyse → evaluate → reflect. For example, begin with 'What do you see?', progress to 'Which visual elements are most important?', then 'How do these elements support your interpretation?', and finally 'Is there another plausible interpretation?' This prevents lessons from remaining at a surface level. (Wang, 2017; Zhao, 2020)

Strengthening Evidence Education

Teachers can require students to use the basic sentence frame 'My view is ... ; the evidence is ... ; because ...' when expressing opinions. After repeated practice, the complexity of the argumentation can gradually be increased.

Facilitating Collaborative Discussion

Group discussion should be prevented from devolving into a simple round of 'everyone says one thing.' Roles such as observer, evidence recorder, questioner, and summariser can be explicitly assigned, so that students undertake different cognitive tasks.

Establishing a Diversified Assessment Mechanism

Assessment should attend not only to final viewpoints but also to the quality of questions raised, use of evidence, logical relationships, attentive listening and responding, and reflective revision. A combination of teacher assessment, peer assessment, and student self-assessment can be employed.

Responsible Use of Digital Resources

Digital images, virtual exhibitions, and online museum collections can help students access a wider range of contemporary artworks, but teachers need to guide students in evaluating information sources, ensuring they do not treat the first interpretation found through an online search as the definitive answer. The value of digital resources lies in supplying material, not in replacing student thinking. (Wang, 2017)

Dimension	Experimental: Pre-test	Experimental: Post-test (Human)	Experimental: Gain	Control: Pre-test	Control: Post-test (Human)	Control: Gain
Problem Awareness	2.00	3.63	+1.63	2.00	2.10	+0.10
Analytical Ability	2.27	3.53	+1.26	2.23	2.50	+0.27
Reasoning Ability	1.83	3.03	+1.20	1.83	2.00	+0.17
Evaluation Ability	1.63	2.73	+1.10	1.67	1.87	+0.20
Reflective Ability	1.40	2.40	+1.00	1.43	1.60	+0.17

Conclusion and Outlook

Using a quasi-experimental design with two parallel Grade 8 classes as the sample, this study empirically examined the effectiveness of contemporary art appreciation teaching in fostering critical thinking, using a five-dimensional critical thinking scale, and employed the Qwen3-8B large language model for auxiliary validation of the scoring data. The findings are as follows: First, contemporary art appreciation teaching has a significant positive effect on the cultivation of critical thinking in Chinese middle school students. The experimental group's post-test mean (15.33) was significantly higher than its pre-test mean (9.13) and the control group's post-test mean (10.07); the mean gain (6.20) was approximately seven times that of the control group (0.90), demonstrating that a structured contemporary art appreciation instructional model can effectively enhance students' abilities in analysis, reasoning, evaluation, and reflection.

Second, evidence awareness and reflective ability are the most prominent dimensions of change. Qualitative data show that experimental-group students' classroom utterances progressed from predominantly 'descriptive layer' contributions to 'reasoning + evaluative layer' and even 'reflective layer' contributions, with some students demonstrating clear metacognitive ability — actively scrutinising their own argumentation processes and adjusting their judgements in light of new evidence.

Third, the large language model (Qwen3-8B) demonstrates preliminary validity as an auxiliary educational assessment tool. Its Pearson correlation with human scores reached 0.876 for the experimental group and 0.959 for the control group; the accuracy of thinking-level judgements on qualitative classroom utterances exceeded 85%, with an agreement rate of 100%. This indicates that large language models hold considerable potential in educational assessment contexts; however, a tendency to underestimate scores in the high-score range (near the maximum; MAE = 1.70) remains, making it necessary to use the model in conjunction with human assessment.

Future research should further expand the sample to include schools in different regions, at different year levels, and of different types, conducting long-term experiments and using validated critical thinking measurement instruments combined with classroom behavioural data. Simultaneously, the refined application of large language models in critical thinking assessment should be further explored, including prompt engineering optimisation, multi-turn conversational scoring, and real-time integration with authentic classroom data. Most importantly, critical thinking should not be simplistically understood as 'daring to contradict the teacher' or 'giving a different answer'; rather, it should be grounded in evidence-based analysis, sound reasoning, careful judgement, and sustained reflection. Only when students gradually develop this thinking habit within the art classroom can the educative value of art education be more fully realised. (Paul & Elder, 2013; Dewey, 1990)

References

- Ministry of Education of the People's Republic of China. (2022). **Compulsory education art curriculum standards (2022 edition)**. Beijing Normal University Press.
- Research Group on Core Competencies. (2016). Core competencies for Chinese student development. **Journal of the Chinese Society of Education, 2016*(10), 1–3.*
- Yin, S. (2015). **Perspectives on core competencies in art.** Hunan Fine Arts Publishing House.
- Teng, S. (2001). **Art and creation.** Shaanxi Normal University Press.
- Efland, A. (2004). **Art and cognition: The role of visual art in education** (X. Li, Trans.). Central Compilation and Translation Press.
- Qian, C. (2013). Contemporary art education: Concepts and practice. **Chinese Art Education, 2013*(4), 12–16.*
- Zhu, H. (2019). The value and pathways of introducing contemporary art into primary and secondary school art classrooms. **Art Education Research, 2019*(11), 112–113.*
- Zhong, Q. (2015). Critical thinking and its teaching. **Global Education Outlook, 44*(1), 34–45.*
- Paul, R., & Elder, L. (2013). **Critical thinking tools** (Y. Hou et al., Trans.). China Machine Press.
- Dewey, J. (1990). **How we think** (W. Jiang, Trans.). People's Education Press.
- Dong, Y., & Dong, C. (2018). Strategies for cultivating critical thinking in middle school students towards core competencies. **Theory and Practice of Education, 38*(29), 56–58.*

- Zhao, X. (2020). Research on teaching strategies for middle school art appreciation classes oriented towards critical thinking cultivation. *Art Teaching, 2020*(5), 45–48.
- Li, M. (2021). Contemporary art appreciation in the context of visual culture and the development of students' higher-order thinking. *Art Education, 2021*(9), 88–91.
- Wang, W. (2017). *An empirical study on cultivating critical thinking in middle school students through "Visual Thinking Strategies" (VTS)* [Master's thesis, East China Normal University].