Vol. 24 (2026): Proceedings of the 2nd International Conference on Linguistics, Psychology and Creative Education (ICLPCE 2026)
E-ISSN: 3105-0522 P-ISSN：3105-0514
DOI: https://doi.org/10.70267/iclpce.202685105

Research on Behavioral Mechanism and System Design of Art Teachers Adopting AI Classroom Assessment Technology Based on Extended TAM

Shiyu Huang*

School of Journalism and Communication, China West Normal University, Nanchong City 637002, China

**Corresponding author: Shiyu Huang.*

Abstract

This study addresses the high subjectivity and heavy workload associated with traditional evaluation in art education, as well as the lack of attention to user adoption behavior in existing AI-in-education research. It aims to reveal the behavioral mechanisms underlying art teachers' sustained use of AI-assisted classroom assessment technologies and to inform system design. An extended Technology Acceptance Model (TAM), incorporating information technology self-efficacy and AU (AI Application Intensity), was constructed and empirically tested using structural equation modeling (SEM) based on questionnaire data (N = 103). The sample primarily consists of art teachers with relatively high experience in using AI tools, and thus the findings are particularly applicable to active technology users. Results indicate that information technology self-efficacy serves as the initial driver of behavior and has the strongest effect on perceived ease of use (β= 0.59). Perceived usefulness significantly influences both BI (Behavioral Intention) (β= 0.39) and application intensity (β= 0.25). A chain mechanism from self-efficacy to ease of use, then to usefulness, intention, and ultimately application intensity is validated. Based on these findings, three design principles are derived, and a teacher-centered conceptual system architecture is proposed. The study provides a coherent linkage between user behavior mechanisms and system design, offering theoretical and practical implications for developing AI-based assessment systems that are usable, adoptable, and sustainable.

Keywords

artificial intelligence-assisted assessment, art education, technology adoption, extended technology acceptance model, system design

1. Introduction

Driven by the dual drive of global education digitization and the national strategy of "artificial intelligence + education", building a data-driven intelligent education system has become a core issue in educational modernization. In this wave of transformation, art education has a particularly urgent need for intelligent upgrading of its evaluation system due to its unique subject attributes.AI-assisted teaching evaluation technology refers to a type of technology or tool that uses artificial intelligence technologies such as computer vision and machine learning to conduct automated, intelligent analysis and feedback on teaching processes and results. Its core function is to assist teachers in completing some evaluation tasks, such as recording and analyzing classroom behavior data, automatically correcting standardized assignments, and generating personalized learning reports [1].

However, traditional art classroom evaluation faces inherent bottlenecks before applying such technologies. First of all, evaluation is highly subjective and relies heavily on teachers' personal experience and aesthetic preferences, making it difficult to objectively and comprehensively reflect students' artistic qualities and creative processes [1]. Secondly, the evaluation tasks are heavy and teachers need to invest a lot of time in reviewing massive visual works, observing classroom behaviors and recording feedback, which takes away the energy for in-depth teaching guidance [2]. Finally, the data generated by art teaching (such as sketches, works, and interactive videos) are mostly unstructured data, which is difficult to systematically analyze and utilize using traditional methods, making it difficult to implement process-oriented and developmental evaluations.

Existing research shows that AI technology provides a new path to break through the above bottlenecks. Early applications focused on processing quantifiable objective questions or structured data [2]. With the advancement of technology, the focus of research is shifting from a single "result evaluation" to a "process evaluation" covering the entire learning cycle such as creative conception, technique application, and cooperation process [3]. AI can provide real-time feedback and multi-dimensional insights by analyzing unstructured data such as classroom behavior and work generation processes, aiming to achieve scientific, personalized and efficient evaluation [1].An international research review also pointed out that AI technology has opened up new opportunities for art education in terms of personalized learning and creativity cultivation, and it is also accompanied by ethical dimensions that need to be carefully considered [4].

Although the technical path represented by AI-assisted teaching evaluation has provided a key solution to solve the long-standing bottleneck problems of art classroom evaluation such as strong subjectivity, heavy teacher load, and difficulty in using process data [1,2], and its practical value has been initially verified - for example, meta-analysis research shows that AI-supported teaching intervention has a positive effect on promoting students' creative thinking [5], and tools such as generative artificial intelligence (GenAI) also bring new possibilities for process evaluation - However, there is still a significant gap in current research. Existing literature mostly focuses on describing the application cases of technical tools and summarizing their teaching effects, mainly answering "what can AI do". However, there is a general lack of in-depth mechanism exploration of the adoption and use behavior of technology users, that is, teachers.

Therefore, the core research question of this article is: What are the core psychological mechanisms and key influencing factors that drive art teachers to continue to adopt and deepen the use of AI assessment tools? In order to answer this question, the research method of this article is: on the basis of systematically sorting out the technical functions and teaching needs, focusing on teachers' behavioral intentions, and constructing an integrated theoretical analysis framework to reveal the internal logic of "why and how teachers continue to use it."

By clarifying the behavioral mechanism of teachers adopting AI assessment tools, this article aims to provide a theoretical basis for the design, promotion and training strategies of related systems, thereby promoting intelligent assessment from "usable" at the technical level to truly "willing to use, easy to use and commonly used" in educational practice.

2. Research Questions

This study aims to go beyond the functional description of AI-assisted assessment technology, to deeply explore the behavioral mechanism of art teachers' adoption and use of this technology, and to provide a theoretical basis for "teacher-centered" system design. To this end, the following three specific research questions (RQ) are proposed:

RQ1: What are the key variables that influence art teachers' adoption of AI-assisted classroom assessment technology?

Existing research mostly focuses on technology functions, but lacks a systematic review of the core factors driving teachers' use. The literature points out that teachers' professionalism and ability are the key, and integrated teaching requires them to have both art professional knowledge and information technology application capabilities [6].

At the same time, teachers' perception of the value of technology, such as whether they believe it can improve the scientificity and accuracy of evaluation [1], is the core cognition that affects its acceptance. Based on this, this study identifies and verifies key variables from the following three dimensions: individual factors (information technology self-efficacy), technology perception factors (perceived usefulness, perceived ease of use) and behavioral variables (willingness to use, practical application intensity), to systematically reveal the behavioral mechanism of art teachers adopting AI assessment tools.

RQ2: Through what paths and mechanisms do the above key factors jointly affect teachers' willingness to use and actual behavior?

Identifying factors is only the first step, and various factors do not have an impact in isolation. For example, teachers' perceptions of technology usefulness may be moderated or mediated by their self-efficacy. However, existing research lacks empirical testing based on theoretical models on the structural relationships and action paths between these factors (which are antecedents and which are mediators). Clarifying this mechanism of action is the key to understanding the logic of teachers' behavioral decision-making. Therefore, this study will construct and verify an integrated theoretical model to reveal the complex mechanism of action from "factors" to "willingness" to "behavior".

RQ3: How to transform the revealed behavioral mechanisms into specific principles guiding the design of "teacher-centered" AI-assisted assessment systems?

The ultimate goal of research is to guide practice and ensure that technical tools are "both feasible and usable." Based on the findings of RQ1 and RQ2, system design should directly respond to teachers' deep needs and behavior-driven logic. For example, if "perceived ease of use" and "result visualization" are proven to be important driving forces for continued use, then the design should prioritize ensuring the simplicity of interaction and transforming AI's data analysis capabilities into intuitive and easy-to-understand teaching insights for teachers (such as automatically generating data reports [1] and supporting diversified evaluations [6]).Therefore, this question aims to translate behavioral theory into actionable design guidelines and principles for system designers.

3. Research Methods

3.1 Theoretical Basis and Model Construction

In order to systematically explore and verify the research questions raised in the previous article, this study is based on the classic technology acceptance model and combines the specific context of art teaching and AI evaluation to construct an extended research model that explains the behavior of art teachers in adopting AI-assisted classroom evaluation.

The technology acceptance model points out that users' willingness to use a new technology is mainly affected by its two core beliefs: perceived usefulness and perceived ease of use [7]. This theory provides a robust foundational framework for understanding teachers' underlying cognitive attitudes toward technological tools. However, the classic TAM model has limitations in explaining specific and complex educational technology adoption behaviors. In the context of art education, AI-driven teaching evaluation is not only the use of a tool, but also involves the in-depth integration of teaching concepts, professional practices and technology. Therefore, it is necessary to introduce key contextual variables and expand the classic model to more accurately capture the multiple factors that influence art teachers' decision-making and their mechanisms of action.

The extended TAM model constructed in this study, on the basis of retaining core variables such as "Perceived Usefulness" (PU), "Perceived Ease of Use" (PEOU) and "Behavioral Intention to Use" (BI), introduces "Information Technology Self-Efficacy (ITSE)" as a key individual trait variable, and "AI Application Intensity" Intensity) as the final behavioral outcome variable. The definitions of each core variable are as follows.

Perceived usefulness refers to the degree to which art teachers believe that using AI-assisted classroom assessment technology can improve the efficiency, effectiveness and scientificity of their teaching assessment work. The core lies in whether teachers recognize that this technology can effectively enhance the objectivity and effectiveness of evaluation [1].

Perceived ease of use means that art teachers believe that the effort required to learn and operate the AI-assisted classroom evaluation system is low, that is, the system is easy to understand and use.

Information technology self-efficacy refers to art teachers' belief in their ability to successfully use information technology (specifically AI evaluation tools) to complete teaching evaluation tasks. This is directly related to whether teachers have the basic knowledge and skill confidence required to control the technology [6].

Intention to use refers to the subjective probability that art teachers intend to use AI-assisted classroom evaluation systems in future teaching.

AI application intensity refers to the depth and breadth of AI evaluation tools actually adopted by art teachers in teaching, including the frequency of use, the diversity of functional uses, and the closeness of integration with teaching processes (such as lesson preparation, classroom interaction, homework evaluation, and reflection). This variable is designed to measure continued and deepening adoption beyond initial adoption.

3.2 Variable Definition and Scale Design

This study uses a structured questionnaire as the main research tool to measure the relevant psychological and behavioral variables of art teachers adopting AI-assisted classroom teaching evaluation technology. The questionnaire was contextualized and revised based on the classic Technology Acceptance Model (TAM) and combined with the art classroom evaluation situation to form a survey tool suitable for the objects of this study.

The questionnaire as a whole consists of three parts. The first part is the basic information of the interviewees, which mainly includes gender, age, education level, teaching period, teaching years, and previous experience in using AI teaching tools, etc., which is used to describe the characteristics of the sample. The second part is the core variable measurement items, which are used to measure the five latent variables of information technology self-efficacy, perceived ease of use, perceived usefulness, intention to use and AU (AI Application Intensity). The third part is open-ended questions, mainly to understand teachers' subjective judgments on the value of AI-assisted art classroom teaching evaluation technology, barriers to use and support needs, so as to provide supplementary reference for result discussion and system design.

Each core variable in this study was scored using a 5-point Likert scale, in which 1 means "strongly disagree", 2 means "strongly agree", 3 means "general", 4 means "somewhat agree", and 5 means "strongly agree". The specific definitions, number of measurement items and theoretical basis of each variable are shown in Table 1. All items were adapted based on existing mature scales, and adjusted in conjunction with the specific context of "AI-assisted art classroom teaching evaluation" to improve the questionnaire's content validity and context adaptability.

Among them, the measurements of perceived usefulness, perceived ease of use, and willingness to use are mainly revised with reference to the classic TAM scale proposed by Davis, which is used to examine teachers' perceptions of the actual value, ease of use, and future use tendencies of AI assessment tools.The measurement of information technology self-efficacy is adapted with reference to related research on computer self-efficacy and is used to assess teachers' confidence in their ability to learn, master, and use AI assessment tools.AU (AI Application Intensity) is measured from the three dimensions of frequency of use, breadth of functional use, and degree of integration with daily evaluation processes, combined with the goals of "continuous use" and "deep application" that this study focuses on, to reflect teachers' integration level of AI evaluation technology in actual teaching.

Before official distribution, the wording of the questionnaire items had been adjusted according to the research topic. For example, the expressions for general information systems in the classic scale are replaced with “AI assessment tools” or “AI-assisted classroom evaluation technology” so that respondents can answer in the specific scenario of primary and secondary art classroom evaluation. At the same time, in order to ensure that the questions are clearly expressed and easy to understand, the entire questionnaire uses concise and direct Chinese expressions, trying to avoid excessive technical terminology burden, so as to improve the accuracy and effectiveness of teachers' filling.

Table 1: Variable definition, number of items and literature sources

Variable	Definition	Number of items	Source	Example Question
Self-efficacy in Information Technology	Teachers' confidence assessment regarding their ability to learn, master, and apply AI assessment technologies.	three	Adapted from the Computer Self-Efficacy Scale by Compeau & Higgins (1995)	Even without external assistance, I am confident I can resolve most operational issues encountered when using AI evaluation tools.
Perceived ease of use	The effort required by teachers to implement AI-assisted classroom assessment technologies.	four	Adapted from the TAM Scale by Davis (1989)	I believe learning to operate AI evaluation tools is straightforward.
Perceived usefulness	Teachers believe that utilizing AI-assisted classroom evaluation technologies can significantly enhance the effectiveness of their teaching assessment practices.	four	Adapted from the TAM Scale by Davis (1989)	Using AI evaluation tools can enhance the efficiency of my teaching assessment tasks (such as grading, recording, and analysis).
Behavioral Intention	The intensity of teachers' tendency to use AI-assisted classroom evaluation technology in future teaching.	three	Adapted from the TAM Scale by Davis (1989)	I am willing to utilize such AI tools in future evaluations of art classroom instruction.
AI Application Intensity	The actual integration level of AI evaluation technologies in teaching by educators encompasses their frequency of use, functional scope, and integration with instructional workflows.	three	This study integrates the 'use' construct from Venkatesh et al. (2011) with the concept of deep application development.	I use AI evaluation tools quite frequently in practice.

According to the final questionnaire design, information technology self-efficacy includes 3 items, perceived ease of use includes 4 items, perceived usefulness includes 4 items, intention to use includes 3 items, and AU (AI Application Intensity) includes 3 items. The measurement quality of this scale tool will be further tested through reliability analysis and confirmatory factor analysis.

3.3 Data Collection and Sample Description

This study used an online questionnaire method to collect data. The questionnaire is distributed to primary and secondary school art teachers and is mainly disseminated through teacher communities, related teaching exchange groups and online education platforms. The survey time is from March to April 2025, and the data collection period is four weeks. In order to ensure data quality, after the questionnaires were collected, samples with incomplete filling, obviously abnormal answering patterns, and signs of invalid responses were eliminated.

In the end, a total of 108 questionnaires were recovered, of which 103 were valid questionnaires, with an effective recovery rate of 95.37%. Judging from the characteristics of the sample, the interviewees are mainly female teachers. They are mainly aged 25-35 years old. Their academic qualifications are mainly undergraduates. Their teaching range covers primary school, junior high school and high school. It should be pointed out that the sample of this study has a relatively high proportion of teachers with high experience in using AI teaching tools. Those who frequently use and are proficient in using AI tools account for 82.52%. Therefore, the research results are more suitable for explaining the adoption mechanism of teachers with certain basic technology exposure.

3.4 Research Hypothesis

Based on literature analysis and theoretical derivation, this study proposes the following hypotheses:

H1: Art teachers’ perceived usefulness has a significant positive impact on their intention to use.

H2: Art teachers' perceived ease of use has a significant positive impact on their intention to use.

H3: Art teachers’ perceived ease of use has a significant positive impact on their perceived usefulness.

H4: Art teachers’ information technology self-efficacy has a significant positive impact on their perceived ease of use.

H5: Art teachers’ information technology self-efficacy has a significant positive impact on their intention to use it.

H6: Art teachers’ willingness to use AI has a significant positive impact on their AU (AI Application Intensity).

H7: Art teachers’ perceived usefulness has a significant positive impact on their AU (AI Application Intensity).

This model assumes that teachers' information technology self-efficacy is an important individual intrinsic factor, which can not only directly affect the intention to use, but also indirectly enhance the intention to use by improving perceived ease of use. Perceived ease of use not only directly promotes BI (Behavioral Intention), but also exerts an indirect effect by strengthening perceived usefulness. Perceived usefulness, as a core cognitive belief, directly drives BI (Behavioral Intention) and has key explanatory power for sustained and in-depth practical application (AI application intensity). Intention to use is a key proximal predictor of behavior, and sustained application intensity also directly depends on whether teachers can continue to perceive the actual value of the tool during use.

This extended model aims to systematically respond to RQ1 and RQ2, and identify the action path and intensity of each influencing factor through empirical testing. The model verification results will provide direct basis for RQ3: A teacher-centered AI evaluation system should be designed to effectively improve teachers’ perceived usefulness and ease of use, and enhance their information technology self-efficacy through training and support, thereby promoting a virtuous cycle from willingness to use to in-depth and continuous application.

3.5 Data Analysis Methods

In order to ensure the rigor and reproducibility of the research conclusions, this study used quantitative analysis methods to process the recovery data and test hypotheses. The specific processes and tools are as follows.

First, SPSS 26.0 software was used for preliminary processing and analysis of all valid questionnaire data. It mainly includes: (1) conducting descriptive statistical analysis on the demographic characteristics of the sample (such as gender, teaching experience, etc.) to present the overall distribution of the sample; (2) conducting a reliability test on the measurement scale of each research variable, and evaluating the internal consistency of the scale by calculating Cronbach’s α; (3) conducting confirmatory factor analysis to test the convergent validity and discriminant validity of the measurement model.

Secondly, in order to test the theoretical model and research hypotheses proposed in this study, a structural equation model was used to conduct path analysis. In view of the sample size (N=103) and model complexity of this study, SmartPLS 4.0 software based on partial least squares (PLS-SEM) was selected for model estimation and testing. The specific analysis steps are as follows:

Model fitness assessment: Focus on testing the quality of the measurement model (external model). The reliability and convergent validity of the scale are evaluated by calculating the composite reliability (CR) and the average variance extracted (AVE); at the same time, the overall goodness of fit of the model is evaluated by calculating the standardized root mean square residual (SRMR). An SRMR value lower than 0.08 is usually regarded as a standard for a good model fit.

Path analysis and hypothesis testing: Run the PLS algorithm to estimate the structural model (internal model), and determine whether each research hypothesis is true by testing the significance of the path coefficient (i.e., p value). This study uses the Bootstrap sampling method (repeated sampling 5000 times) to obtain the t statistic and confidence interval of the path coefficient. In addition, the overall explanatory strength of the independent variables on the dependent variable is evaluated by calculating the coefficient of determination (R^2) of the endogenous latent variables. The significance level for path coefficients was set at $p < 0.05$.

Mediation effect analysis: For the potential mediating paths in the model (such as "information technology self-efficacy → perceived ease of use → perceived usefulness"), the Bootstrap method (5000 sampling times) was also used to test. If the 95% bias-corrected confidence interval of the indirect effect does not include 0, it indicates that the mediation effect is significant.

Through the above systematic and transparent data analysis process, we aim to ensure the reliability and scientificity of the research conclusions and provide a solid empirical basis for subsequent discussions and suggestions.

4. Empirical Analysis and Results

In order to verify the theoretical model and research hypotheses constructed in Chapter 3 of this article, and to reveal the behavioral mechanisms that influence art teachers to adopt AI-assisted teaching evaluation tools, this study is based on the data collected from a special questionnaire survey and uses structural equation modeling (SEM) and other methods for empirical testing.

4.1 Research Objects

This study distributed questionnaires through online channels, and a total of 108 questionnaires were collected. After eliminating invalid questionnaires such as incomplete and regular answers, 103 valid questionnaires were obtained, with an effective recovery rate of 95.37%.

Analysis of 103 valid respondents showed that female teachers accounted for the vast majority (84.47%). The age distribution is dominated by young and middle-aged teachers, with the highest proportion (36.89%) aged 26-35, and 31.07% aged 25 and under. The educational background is mainly bachelor degree (64.08%). The distribution of teaching sections is relatively balanced, with the proportion of teachers in primary school, junior high school and high school being 33.98%, 31.07% and 34.95% respectively. In terms of teaching experience, teachers with 4-10 years accounted for the highest proportion (38.83%).

In terms of technical background, a notable feature is that the interviewed teachers are generally familiar with AI teaching tools: the total proportion of teachers who self-assessed that they "frequently use" and "proficiently use" AI-assisted teaching tools is as high as 82.52% (43.69% and 38.83% respectively), while those who "never use" account for only 1.94%. This shows that the sample of this study mainly comes from a group of teachers who have a certain level of attention and experience in using AI technology. Therefore, the conclusion of this study focuses more on revealing the behavioral mechanism of teachers who already have a good technical foundation in adopting special assessment tools. Caution should be exercised when extending it to all teachers (especially technical beginners). This sample also provides a corresponding data basis for understanding the deepening adoption mechanism of this specific group.

4.2 Measurement Model Testing

First, the reliability and validity of the measurement scales of the five core latent variables in the model (information technology self-efficacy ITSE, perceived ease of use PEOU, perceived usefulness PU, intention to use BI, and AU (AI Application Intensity) AU) were tested, as shown in Table 2.The reliability analysis results show that the Cronbach's α coefficient of each variable ranges from 0.88 to 0.92, and the combined reliability (CR) values are all greater than 0.85, indicating that the scale has good internal consistency reliability. Secondly, the validity was tested through confirmatory factor analysis (CFA). The standardized factor loadings of all measurement items were greater than 0.72, and the average variance extracted (AVE) values were greater than 0.60, proving good convergent validity. The square root of each variable's AVE is greater than the correlation coefficient between it and other variables, indicating good discriminant validity.

The specific discriminant validity test results (Fornell-Larcker Criterion) are shown in Table 3. The bolded diagonal values in the table are the square roots of AVE of each variable, and the off-diagonal lines are the correlation coefficients between variables. It can be seen that the square root of each latent variable AVE is significantly larger than its correlation coefficient with other latent variables, which further confirms that the measurement model has good discriminant validity. The measurement model fit was assessed using the standardized root mean square residual (SRMR = 0.044), which is below the recommended threshold of 0.08, indicating good model fit.

Table 2: Results of reliability, validity, and confirmatory factor analysis for the measurement model

Latent Variable	Measurement Item (Example)	Standardized Factor Loadings	Cronbach 's α	Combined Reliability (CR)	Mean variance estimator (AVE)
Information Technology Self-Efficacy (ITSE))	ITSE1: I am confident in learning and utilizing new information technologies (including AI tools).	0.85	0.90	0.92	0.65
	ITSE2: Even without external assistance, I am confident I can resolve most operational issues encountered when using AI evaluation tools.	0.88			
	ITSE3: I am confident in my ability to effectively integrate AI assessment tools into my art teaching evaluation practices.	0.82			
Perceived ease of use (PEOU)	PEOU1: I believe learning to operate AI evaluation tools is straightforward.	0.78	0.88	0.89	0.62
	PEOU2: I can effortlessly and proficiently use AI evaluation tools to complete the assessments I need to conduct.	0.83			
	PEOU3: My interaction with the AI evaluation tool is clear, straightforward, and easy to understand.	0.79			
	PEOU4: Overall, I find the AI evaluation tool easy to use.	0.80			
Perceived Utility (PU)	PU1: Using AI evaluation tools enhances the efficiency of my teaching assessment tasks (such as grading, recording, and analysis).	0.86	0.92	0.93	0.68
	PU2: Using AI evaluation tools enhances the objectivity and scientific rigor of my teaching assessments.	0.88			
	PU3: The AI assessment tool provides me with valuable insights, helping me better understand students' learning progress.	0.82			
	PU4: Overall, I find the AI assessment tool to be useful for evaluating my art teaching practices.	0.85			
Behavioral	BI1: I am willing to use such	0.90	0.91	0.92	0.70

Latent Variable	Measurement Item (Example)	Standardized Factor Loadings	Cronbach 's α	Combined Reliability (CR)	Mean variance estimator (AVE)
Intention (BI)	AI tools in future evaluations of art classroom instruction.				
	BI2: If given the opportunity, I plan to frequently use AI evaluation tools.	0.87			
	BI3: I would recommend using the AI evaluation tool to my colleagues.	0.84			
AI Application Intensity (AU)	AU1: I use AI evaluation tools quite frequently in practice.	0.75	0.89	0.90	0.61
	AU2: I utilized multiple features of the AI assessment tool to support teaching evaluations at various stages.	0.82			
	AU3: The AI assessment tool has been seamlessly integrated into my daily art teaching evaluation process.	0.80			

Table 3: Discriminant validity test results (Fornell-Larcker Criterion)

Variable	1	2	3	4	5
1. ITSE	0.806				
2. PEOU	0.59**	0.787			
3. PU	0.45**	0.44**	0.825		
4. BI	0.48**	0.52**	0.54**	0.837	
5. AU	0.40**	0.42**	0.50**	0.53**	0.781

4.3 Descriptive Statistics and Correlation Analysis

The mean, standard deviation and Pearson correlation coefficient of each core variable are shown in Table 4 (*p<0.05, **p<0.01). The means of information technology self-efficacy (M=3.21), perceived ease of use (M=3.23), perceived usefulness (M=3.18), intention to use (M=3.21) and AU (AI Application Intensity) (M=3.20) are between 3.18 and 3.23 on the 5-point scale, indicating that teachers' overall attitude towards each variable is neutral and slightly positive. Different from the "high intention, low usage" phenomenon that often occurs in previous studies, the mean value of AU (AI Application Intensity) (M=3.20) and BI (Behavioral Intention) (M=3.21) in this sample are similar, and there is no significant gap. Correlation analysis shows that, as shown in the matrix in the lower half of Table 4, there is a significant positive correlation between all variables ($p < 0.01$), which provides preliminary support for the path relationship hypothesized in the structural model.

Table 4: Descriptive statistics and correlation coefficient matrix between variables (N=103)

Variable	mean M	standard deviation SD	1	2	3	4	5
1. Self-efficacy in Information Technology	3.21	0.65	—				
2. Perceived ease of use	3.23	0.62	.59**	—			
3. Perceived usefulness	3.18	0.68	.45**	.44**	—		
4. BI (Behavioral Intention)	3.21	0.64	.48**	.52**	.54**	—	
5.AU (AI Application Intensity)	3.20	0.66	.40**	.42**	.50**	.53**	—

*Note: ** indicates p <0.01 (two-tailed test).*

4.4 Structural Model and Hypothesis Testing (Path Analysis)

Structural equation modeling was used to test the theoretical model and research hypotheses, and The structural model fit was evaluated using PLS-SEM. The standardized root mean square residual (SRMR = 0.059) fell within the acceptable range (< 0.08), indicating satisfactory model fit. The path coefficients and hypothesis test results are summarized in Table 5.

Table 5: Results of Path Analysis and Hypothesis Testing for the Structural Model

path relationship	hypothesis	Standardized path coefficient (β)	Standard error (S.E.)	t value	p value	result (of inspection)
Perceived usefulness → Intention to use	H1	0.39	0.08	4.875	<0.001	support
Perceived ease of use → Intention to use	H2	0.30	0.07	4.286	<0.001	support
Perceived ease of use → Perceived usefulness	H3	0.44	0.09	4.889	<0.001	support
Information technology self-efficacy → Perceived ease of use	H4	0.59	0.10	5.900	<0.001	support
Information technology self-efficacy → BI (Behavioral Intention)	H5	0.19	0.09	2.111	<0.05	support
BI (Behavioral Intention) →AU (AI Application Intensity)	H6	0.37	0.08	4.625	<0.001	support
Perceived usefulness →AU (AI Application Intensity)	H7	0.25	0.09	2.778	<0.01	support

The direct effect test results show that all seven proposed research hypotheses (H1-H7) are supported. Specifically:

Perceived usefulness has a significant positive impact on BI (Behavioral Intention) (β=0.39, p<0.001) and AU (AI Application Intensity) (β=0.25, p<0.01). H1 and H7 are established.

Perceived ease of use significantly positively affects BI (Behavioral Intention) (β=0.30, p<0.001) and perceived usefulness (β=0.44, p<0.001), H2 and H3 are established.

Information technology self-efficacy not only directly and positively affects BI (Behavioral Intention) (β=0.19, p<0.05, H5 is established), but also has an extremely strong positive impact on perceived ease of use (β=0.59, p<0.001), H4 is established.

In this research model, BI (Behavioral Intention) is the most critical direct factor in predicting AU (AI Application Intensity) (β=0.37, p<0.001), and H6 is established.

4.5 Analysis of the Mediating Effect

In order to further explore the complex interaction mechanism between key variables in the model, especially the indirect influence paths of variables such as information technology self-efficacy, perceived ease of use, and perceived usefulness, this study conducted a Bootstrap method (repeated sampling 5,000 times) test and decomposition of the total effect, direct effect, and indirect effect based on the verified path coefficients of the structural model. Set the direct paths of “information technology self-efficacy→AI application intensity” and “perceived ease of use→AI application intensity”. The effect decomposition results are detailed in Table 6.

Table 6: Effect Decomposition and Mediation Effect Analysis Among Key Variables

Effect Path	Direct effect	Indirect effect	Total effect	Proportion of indirect effect	Mediation type
Information technology self-efficacy → Perceived usefulness	0.19*	0.26**	0.45**	57.8%	partial mediation
Information technology self-efficacy →AU (AI Application Intensity)	n.s. (Not specified)	0.20**	0.20**	100%	full mediation
Perceived ease of use →AU (AI Application Intensity)	n.s. (Not specified)	0.22**	0.22**	100%	full mediation
Perceived usefulness →AU (AI Application Intensity)	0.25**	n.s.	0.25**	0%	indirect-only effect
BI (Behavioral Intention) →AU (AI Application Intensity)	0.37**	n.s.	0.37**	0%	indirect-only effect

The core mediating role of information technology self-efficacy: The total effect of information technology self-efficacy (ITSE) on perceived usefulness (PU) is 0.45, of which the indirect effect (by improving perceived ease of use) accounts for 57.8%. At the same time, its direct effect on the final behavioral outcome "AI application intensity" is not significant, and the total effect of 0.20 is entirely contributed by the indirect effect (through the mediator of perceived ease of use, perceived usefulness and intention to use). This highlights its role as a basic and initiating variable in the driving behavior chain, and its impact must be achieved by improving teachers' cognition and attitude towards the system.

Indirect driving effect of perceived ease of use: Analysis shows that perceived ease of use has no significant direct effect on AU (AI Application Intensity), and its total effect of 0.22 is completely realized through the indirect path (that is, by enhancing perceived usefulness and willingness to use). This shows that the "ease of use" of the system itself does not directly lead to in-depth use behavior. Its core value is to help teachers establish a positive perception of the "usefulness" of the tool and form a willingness to use it by lowering the threshold for use, thereby indirectly promoting the deepening of application.

The direct driving effects of perceived usefulness and intention to use: The effects of perceived usefulness and intention to use on AU (AI Application Intensity) are significant direct effects, and there is no significant indirect path. This suggests that once teachers are convinced that the tool is valuable (perceived usefulness) and form an intention to use it, this belief and intention will directly translate into deeper and more extensive actual use behavior.

To sum up, the mediation effect analysis clearly reveals the hierarchical nature of the behavior driving mechanism: information technology self-efficacy is a deep-seated initiating factor that works entirely through mediation; perceived ease of use is a key transmission hub, and its value lies in empowerment rather than direct driving; and perceived usefulness and willingness to use are direct "driving engines" close to the behavioral end. This finding provides a more refined theoretical basis for precise, hierarchical intervention strategies (such as prioritizing training to enhance teachers' sense of efficacy to activate the entire influence chain).

4.6 Explanation and Discussion of Behavioral Mechanism

The results of the empirical analysis systematically reveal the behavioral mechanism of adopting AI-assisted assessment tools among the specific sample of this study (i.e., the group of teachers with higher exposure to AI technology).

This study verified a clear chain-driven path. Teachers' information technology self-efficacy (ITSE) is the key psychological initiating factor in the entire behavioral process, which can significantly and powerfully enhance their perception of the "ease of use" of the tool (β=0.59). This perception of ease of use has a dual role: on the one hand, it directly promotes "intention to use"; on the other hand, it indirectly promotes the formation of intention by enhancing the perception of "usefulness". In the end, strong willingness to use and continuous perception of usefulness jointly determine the actual and in-depth "AI application intensity." The phenomenon that the mean levels of "AI application intensity" and "intention to use" are similar observed in this sample may be related to the fact that the sample teachers themselves have a higher level of technical readiness and are therefore able to more smoothly transform their intentions into sustained actions.

The "efficacy-ease of use-usefulness-behavior" driving chain revealed in this study is consistent with the core logic of the classic Technology Acceptance Model (TAM) [7], and has been further confirmed in educational technology application scenarios. For example, research in the field of language education points out that the effective integration of AI tools relies heavily on teachers' clear understanding of their teaching value and ease-of-use experience in specific scenarios [8]. This jointly shows that converting technical capabilities into user-perceivable value ("value explicitness") and providing support to adapt to different capability stages ("supporting hierarchical advancement") are general principles for promoting in-depth adoption. This concept has also been confirmed in other aspects of AI education applications. For example, in the design of AI courses for students, research also emphasizes the need to adopt inclusive and hierarchical strategies to adapt to the diverse starting points of learners [9].

The above mechanism provides a direct basis for the design of "teacher-centered" AI assessment system. First of all, in view of the strong influence of the "self-efficacy → ease of use" path (β=0.59), the system design must be committed to giving priority to reducing the cognitive load and technical anxiety in the early stage of use through minimalist interaction logic, clear guidance and immediate positive feedback for operation, so as to quickly establish a sense of control for teachers (especially those with low confidence). Secondly, the system must clearly and intuitively transform AI's data processing and analysis capabilities into teaching value that teachers can perceive (such as significantly improving evaluation efficiency and providing previously difficult-to-obtain learning insights) to continuously enhance its perceived usefulness, which is the core driving force for continued use. Finally, it is necessary to design hierarchical and advanced support functions and resources to adapt to the needs of teachers with different experience levels, from novices to experienced ones, to support them from initial attempts to deepening integration, and ultimately to bridge the gap between "use" and "in-depth application".

In summary, the behavioral mechanism empirically revealed in this study provides clear engineering requirements guidance for the subsequent design of "teacher-centered" AI-assisted assessment systems. Specifically, the verified user behavior model can be directly transformed into three core system requirements: First, in response to the key role of "information technology self-efficacy" (ITSE) as the starting point of behavior, the system needs to integrate low learning threshold guidance and immediate assistance mechanisms to quickly establish and enhance teachers' confidence in use. Secondly, in order to meet the requirement of "perceived usefulness" (PU) as the core driving force, the system must have the ability to visualize the analysis results and make decision-making outputs, ensuring that the underlying data analysis of AI is transformed into insights that teachers can intuitively understand and be directly used for teaching decision-making. Finally, in order to ensure a powerful transformation path from "intention to use" (BI) to deep "AI application intensity" (AU), the system design must support seamless embedding with the existing teaching process and provide a feedback and incentive closed loop that promotes continued use. This mapping from behavioral mechanisms to engineering requirements clearly builds a bridge from user research to system engineering practice, enabling system design to accurately respond to teachers' deep psychological and behavioral logic.

5. From Model to System Design

Based on the aforementioned theoretical construction and empirical testing, this study not only reveals the psychological behavioral mechanism of art teachers adopting AI evaluation tools, but also provides direct empirical basis for "teacher-centered" system design. This chapter aims to complete the engineering transformation from "behavioral model" to "design principles" to "system conceptual framework", and provide specific guidance for subsequent development.

5.1 Extraction of Design Principles Based on Empirical Conclusions

Based on the empirical findings from Chapter 4, the system design should focus on three key areas: enhancing self-efficacy, reducing perceived costs, and strengthening value perception.

Principle 1: Minimize Initial Cognitive Load

Empirical Basis: Teachers' information technology self-efficacy is the strongest predictor of perceived ease of use ($\beta = 0.59$). Reducing complexity directly strengthens confidence and adoption.

Design Implication: Implement one-click core workflows (e.g., one-click batch review), replace nested menus with graphical/contextual guidance, and minimize steps for key evaluation tasks. Provide immediate feedback to reinforce perceived control for first-time users to establish a sense of competence.

Principle 2: Make Pedagogical Value Explicit

Empirical Basis: Perceived usefulness strongly influences willingness to use ($\beta = 0.39$) and depth of use ($\beta = 0.25$), yet current perception remains moderate ($M = 3.18$), indicating unmet potential.

Design Implication: Translate AI outputs into ready-to-use teaching artifacts—e.g., auto-generated standardized academic reports for parents/portfolios, and visual analytics on student ability growth (e.g., composition, color use). Each feature must pass the "time saved or insight gained" test versus manual workflows.

Principle 3: Support Progressive and Differentiated Adoption

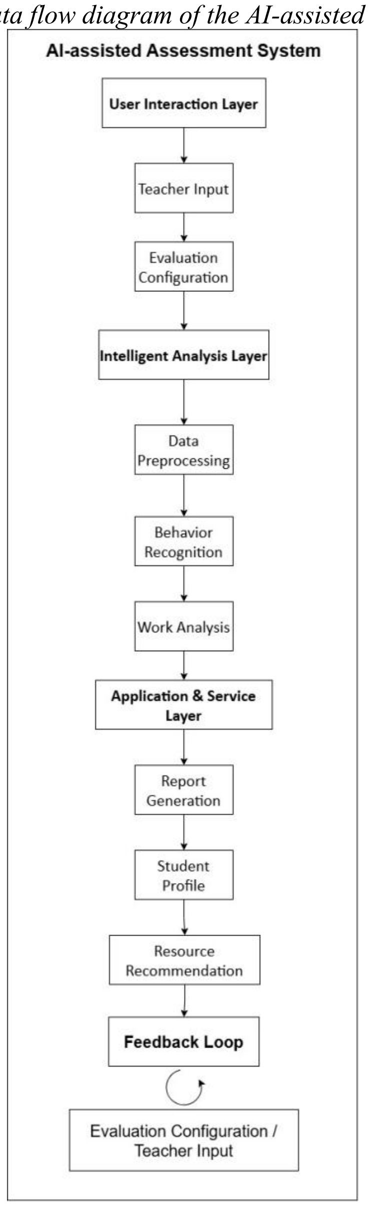
Empirical Basis: Aligns with constructivist-oriented generative AI trends in art education, emphasizing transition from tool user to instructional designer through scaffolded integration. This is consistent with the research trend of generative AI in the field of art education, which advocates the transformation of teaching models based on constructivist theory by supporting the deep integration of the creative process, reflection and collaboration [10].

Design Implication: Offer tiered support—preset templates and automated reports for novices; customizable metrics and deep analytics for experts. Include adaptive guidance (step-by-step vs. shortcuts) and foster peer learning via communities and shared case libraries to grow with teachers' expertise.

6. System framework

Based on the behavioral mechanisms and design principles revealed by empirical research, this study constructed a conceptual framework for a teacher-centered AI-assisted art classroom teaching evaluation system. Its three-level architecture and core data flow are shown in Figure 1. This framework clarifies the complete engineering logic from multi-modal data input, intelligent analysis and processing, to finally generating output with perceptible teaching value, and forming a feedback optimization closed loop.

Figure 1: Three-layer architecture and data flow diagram of the AI-assisted art classroom teaching evaluation system



6.1 Overall System Data Flow and Workflow

The system is built around a closed loop of "data collection → intelligent analysis → value output → feedback optimization". The process starts with multi-modal data collection. The system automatically or the teacher enters raw data such as students' digital paintings, classroom process videos, and teachers' verbal comments. Subsequently, the data preprocessing module performs standardized cleaning, format conversion and key information extraction on the original data (such as segmenting individual student behavior segments from videos). The processed structured data is sent to the intelligent analysis module, which is the "brain" of the system and is embedded with AI models such as computer vision and natural language processing. It is responsible for multi-dimensional analysis of composition, color, technique and other aspects of the work, as well as identification of participation and collaboration in classroom behaviors. The underlying characteristics and insights generated by the analysis are integrated by the evaluation generation module, which automatically generates structured evaluation reports, capability radar charts and personalized improvement suggestions based on preset or teacher-defined evaluation rules. Finally, all outputs are presented to teachers through the teacher interaction and feedback module. Teachers can review and revise the evaluation results and feed them back to students. At the same time, the teacher's adoption, modification or rejection of the system's suggestions will flow into the system as new feedback data for optimizing the AI model, thereby achieving closed-loop iteration. This framework is a conceptual architecture and not a deployed prototype system. This closed-loop design concept aims to respond to the recent call for AI education tools to go beyond a single function and move toward deep integration of multiple roles such as "evaluation and feedback" and "intelligent guidance" [11]. At the same time, it also recognizes that, as the experience of promoting learning analytics technology shows, the effectiveness and sustainability of technical tools are not only related to their own design, but also deeply dependent on continuous interaction with teaching practices, a "feedback optimization" mechanism based on data iteration, and broader institutional data culture and support strategies [12].

6.2 Three-level Framework and Modular Input and Output

To implement the aforementioned workflow, the system is logically divided into the interaction layer, analysis layer, and application layer. The input/output specifications for the core modules of each layer are detailed in Table 7.

Table 7: Input/Output Definitions for Core Modules of the Three-Level System Framework

Hierarchy	Core module	Main input	Algorithm/Function Description	Main output	Service object/purpose
Alternation of bed	Evaluation Configuration Module	The evaluation objectives, criteria, and types of classroom/assignment work set by the teacher	Provide a drag-and-drop rule configuration interface; support quick import and fine-tuning from preset template libraries (e.g., "Primary School Watercolor Painting Evaluation").	Structured evaluation task instructions	Teacher (Task Configuration)
Analysis Layer	Result Visualization Module	The raw data reports, charts, and labels generated by the analysis layer	An integrated data visualization library that automatically generates capability radar charts, classroom engagement time-series charts, and visual feature heatmaps of projects; supports interactive filtering of report elements.	An interactive evaluation report rich in both text and images	Teachers (making instructional decisions), students (receiving feedback)
	Interactive Feedback Module	Teacher's supplementary evaluations, students' questions, and	It features a lightweight comment system that allows users to like, edit, and supplement their comments, and records	Structured feedback records, conversation logs	Teacher-student interaction and teaching reflection

Hierarchy	Core module	Main input	Algorithm/Function Description	Main output	Service object/purpose
		system-generated formative suggestions	teacher-student conversation logs.		
	Multimodal Data Preprocessing Module	Classroom videos, student work images, audio, and text logs	Video processing: Utilizes face/pose detection for student localization and behavior segment segmentation. Image processing: Performs standardization operations such as perspective transformation and color space conversion. Audio/text processing: Implements speech activity detection (VAD), automatic speech recognition (ASR), and text cleaning.	Standardized, structured multimodal data	Provide high-quality input for the upper-level analysis module
Application and Service Layer	Classroom Behavior Recognition Module	The preprocessed video and audio streams	Behavior recognition: Recognize gestures such as “raise hands” and “discuss” based on gesture estimation models; evaluate speeches through sound source localization and voice emotion analysis. Participation calculation: Comprehensive visual attention and voice activity to generate participation index.	Behavioral sequence tags, engagement index, and abnormal behavior alerts	Provide a quantitative basis for process evaluation
	Intelligent Work Analysis Module	Student work images, creation process sketches, and text descriptions	Visual Analysis: Extract color histograms and color harmony; utilize a pre-trained CNN model for composition classification and technique identification (e.g., “wet painting technique”). Text Analysis: Perform keyword extraction and sentiment analysis on the creative descriptions.	Multi-dimensional work scoring, visual feature tags, and creative descriptions	Conduct a multi-dimensional and relatively objective analysis of the work
	Student Learning Report Generation Module	Behavior data, work data, student history data, evaluation criteria	Report Generation: Utilizing a template engine and natural language generation (NLG) technology, structured data is combined into a coherent narrative report. Attribution Analysis: Simple statistical	Structural Chemistry Report, Personalized Learning Path Recommendations	Support teachers in implementing targeted instructional interventions

Hierarchy	Core module	Main input	Algorithm/Function Description	Main output	Service object/purpose
			methods are applied to identify capability gaps.		
	Student Digital Profile Service	The complete lifecycle behavior and outcome data generated by students within the system	Continuously aggregate students' data across knowledge and skills, process performance, and emotional attitudes to establish a dynamically updated multi-dimensional tagging system and growth trajectory.	Student Digital Profile Model, Vertical Growth Trajectory Chart	Support long-term developmental assessment and personalized resource recommendations
	Teaching Resource Recommendation Service	Student Learning Report, Digital Profile, Teaching Resource Library, Curriculum Standards	Collaborative filtering: Recommendations based on students' historical resource usage patterns. Rule matching: Precisely link technical micro-lessons and case studies to specific suggestions (e.g., “Enhance color contrast”).	Personalized list of teaching/learning resources	Teachers (lesson preparation support), Students (self-directed learning)
	Manage Data Interface Services	Aggregated class/grade evaluation data and teaching analysis reports	Provide standardized RESTful API interfaces for university or regional education management platforms, delivering anonymized macro-level teaching quality data in accordance with standards such as xAPI.	Standardized data packages, macro quality analysis reports	School administrators, education researchers

6.3 System Application Case Demonstration

The following simplified case illustrates the system's workflow from data input to value output. This example demonstrates the system's operational process and does not represent actual system output. Case scenario: During a watercolor class titled “The Colors of Summer,” Teacher Zhang utilized this system to assist in their assessment.

After the course concluded, Teacher Zhang used the system's interactive terminal to batch-upload the digital watercolor artworks of all students (e.g., Student A's “Summer Lotus Pond”.jpg) and authorized the system to analyze the video recording of the class session. System Processing and Output: In the data preprocessing stage, the system automatically corrects the artwork and extracts close-up segments of each student's creative work from the video.

Intelligent analysis: The visual analysis engine of the painting “Summer Lotus Pond” identified that it adopts “central composition”. Color analysis shows that green color accounts for 60% and the saturation is harmonious. Brushstroke analysis detects the use of “wet painting method”. The behavior analysis engine identified from the video that the student was highly focused when creating and had brief interactions with classmates sitting next to him. Based on the above information, the evaluation generation engine calls the “watercolor painting” evaluation rule to generate a structured report.

Teachers gain value: The system can output results in a short processing time. Teacher Zhang saw the following output on the interactive terminal: In terms of quantitative scoring, the composition is 85/100; the color matching is 90/100; the qualitative comments are that the theme of the work is prominent, the colors are fresh and harmonious, and it better expresses the summer atmosphere. The use of wet painting method makes the color transition natural; for personalization suggestions, you can try to add a little warm color (such as light orange) in the lower right corner of the picture to enhance color contrast and spatial hierarchy.

You can refer to the light and shadow processing in Monet's “Water Lilies” series; at the classroom insight level, the system also provides an overall classroom participation report, reminding Teacher Zhang that two students have been offline for a long time during the creative process.

Results: The system generated subject comments and suggestions based on the preset classroom evaluation dimensions and image analysis results. Teacher Zhang only fine-tuned individual expressions and quickly completed an in-depth evaluation of the student's work. The system automatically archives this evaluation to the student's digital growth file, and associates the corresponding technique micro-course for Student A in the resource library based on the suggestion point of “color contrast”.

To clearly illustrate how the system transforms low-level data features into high-level instructional value, the table below (Table 8) summarizes the core mapping relationships in this case study:

Table 8: Example of core system rule mapping in case scenarios

Input Feature	Analysis and Processing Module	Evaluative dimension	Final output format
The painting contains 60% green tones with high saturation.	Image Analysis Engine (Color Analysis Submodule)	Color Coordination and Harmony	Quantitative score (90/100), qualitative comment (“The colors are fresh and harmonious”)
Main layout structure of the image	Image Analysis Engine (Composition Analysis Submodule)	Rationality of composition	Quantitative score (85/100)
The “wet brushstroke” feature was detected.	Image Analysis Engine (Stroke/Technique Recognition Submodule)	The appropriateness of the technique used	Qualitative comment (“The use of the wet painting technique ensures a natural transition”)
The “high concentration” and “communication” behaviors identified in the video	Classroom Behavior Recognition Engine	The level of investment in learning and the degree of cooperation	Integrate comprehensive evaluations to create a student's digital profile
Total offline time statistics (for multiple students)	Classroom Behavior Recognition Engine (Aggregated Analysis)	Overall classroom engagement	Independent classroom insight reports and early warnings

This framework demonstrates, through clear data flows, modular functional design, and real-world teaching scenarios, how findings from behavioral mechanism research—such as enhancing usability, clarifying value, and supporting hierarchical structures—can be translated into concrete engineering implementation pathways, providing a actionable blueprint for subsequent system development.

6.4 Scenario-based Simulation and Verification of the Conceptual Framework

To demonstrate the engineering feasibility of the aforementioned conceptual framework and its responsiveness to the design principles proposed in empirical research, this section simulates and validates the framework's core workflow using an integrated typical teaching scenario. This validation aims to illustrate (rather than execute) how the framework transforms multimodal inputs into value outputs that directly address teachers' core needs, thereby providing engineering-level support for the three design principles: minimizing initial cognitive load, explicitizing teaching value, and facilitating tiered and progressive adoption.

Simulation scenario: Taking the primary school art lesson “The Artistic Conception of Ink Painting” as an example, Teacher Li must complete two core assessment tasks: “Evaluating Students' Ink Painting Projects” and “Observing the Classroom Creative Process.”

1. Input Data:

Work Images: Digital ink painting assignments by all students (e.g., Student B's “Distant Mountains,”.png). Process Data: Video clips from the lesson recording capturing students' painting process. Rule Input: Through the “Evaluation Configuration Module” in the interactive interface, Teacher Li selected the “Primary School Ink Painting” evaluation template (which includes dimensions such as composition, brushwork, and artistic conception) from the preset library as the basis for this analysis.

2. The processing workflow and module coordination system operates sequentially according to the workflow of “data collection → intelligent analysis → value generation → feedback-driven optimization” and the module functions defined in Table 7:

Step 1 (Data Preprocessing): The “Multimodal Data Preprocessing Module” in the analysis layer automatically performs image normalization on the “Far Mountains”.png file and extracts key student B's drawing activity segments from the classroom video.

Step 2 (Intelligent Analysis): Work Analysis: The “Intelligent Work Analysis Module” utilizes a pre-trained computer vision model to extract visual features from the artwork, such as identifying that “the blank space accounts for approximately 40%,” “the ink gradation consists of five levels,” and “the primary technique employed is the textural brushstroke method.” Behavior Analysis: The “Classroom Behavior Recognition Module” analyzes the video footage and identifies that “the artist focused on painting continuously for about 12 minutes,” with three instances of dipping and mixing ink observed during this period.

Step 3 (Evaluation Generation and Value Packaging): The “Student Learning Report Generation Module” in the Application and Service Layer receives the aforementioned characteristics and identification results, and performs mapping and synthesis based on the evaluation criteria for “Primary School Ink Painting.” For example, “a high proportion of negative space” is interpreted as “strong compositional awareness, understanding the principle of treating negative space as negative area”; “rich ink tonal gradations” are linked to “the application of texturing techniques,” indicating “basic mastery of ink wash variations in dryness, wetness, and density”; and “ focused behavior” serves as a positive criterion for formative assessment.

3. Output and Value Presentation

Within minutes, the system provides Teacher Li with the following structured output through the interaction layer:

Decision-making quantitative and qualitative report: Quantitative score: composition: 88/100; pen and ink technique: 85/100; artistic conception expression: 82/100. Qualitative comments: “The student can actively use white space to create a sense of space. The ink color changes are relatively rich, and the overall picture has begun to have the charm of ink and wash. He shows a high degree of concentration during the creation process.” Personalized improvement suggestions: “You can try to apply a light ocher color on the outline of the distant mountains to enhance the texture of the mountain and the subtle changes in color of the picture (see the 'Light Crimson Landscape' technique micro-lesson in the resource library).”

Visual insights: Simultaneously generate a “radar chart of overall writing and ink usage levels in the class” and a “student individual growth trend chart in various dimensions”.

Hierarchical resource recommendation: Based on the suggestion of “Qianjiang Landscape”, the system automatically pushes relevant masterpiece appreciation pictures and technique explanation video links in the sidebar.

4. Evaluation of the achievement level of design objectives

This simulation verification shows that the engineering logic of the conceptual framework can effectively support the three core design principles extracted in Section 5.1:

Supports “minimizing initial cognitive load” (response to Principle 1): Teacher Li only needs two steps of “uploading works/videos” and “selecting an evaluation template” to start the complex evaluation process. This is in line with the conclusions of technology adoption research on teachers in specific subjects (such as mathematics teachers), which emphasizes the importance of simplifying the initial operation process and lowering the threshold for technology use to improve teachers' (especially teachers with clear task goals) adoption intention [13]. The system automatically handles the entire process from feature extraction to report generation, transforming the strong driving path of “information technology self-efficacy → perceived ease of use” ($\beta = 0.59$) revealed in empirical research into an engineering implementation of “minimalist operation → instant, complete output”, directly helping teachers establish an initial sense of control.

Achieving “the explicitness of teaching value” (response to Principle 2): The system output is not raw data or technical labels, but comments, scores and specific suggestions that are directly embedded in the teaching context and can be used directly by teachers or slightly adjusted. This directly responds to the empirical finding that teachers’ “perceived usefulness” is the core factor driving their intention to use and continued use behavior [14].

Reflecting “layered and progressive support” (response to Principle 3): The framework not only provides Teacher Li with a standardized evaluation template (lowering the threshold for novices), but also allows him to customize the evaluation dimensions and weights in the “Evaluation Configuration Module” in the future (supporting in-depth use by experts). At the same time, the “Case Library” and “Resource Push” functions provided by the system provide a supporting path for teachers to evolve from tool users to instructional designers, which helps promote the continuous transformation from initial intention to use to in-depth application intensity (β=0.37). This flexible design that supports the development of teachers from novices to experts is consistent with the broader AI education innovation research perspective that emphasizes the integration of multiple factors [15]. It not only serves the ability growth of individual teachers, but its endogenous high-quality assessment data flow also provides the underlying data interface and practical paradigm for the future construction of AI-empowered curriculum and project design frameworks at the school level [16].

Through the above scenario-based simulation verification, the conceptual system framework proposed in this study demonstrates its complete engineering logic for transforming multi-modal data into actionable teaching insights. This framework does not stop at the list of functions, but through clear module division of labor and data flow design, it empirically demonstrates how to transform the key driving factors in the user behavior model (self-efficacy, perceived usefulness, intention to use) into specific and achievable system characteristics, thus providing a logically verified blueprint for the subsequent development of prototype systems.

7. Conclusion

At the theoretical level, the main contribution of this study is to systematically integrate the key individual trait variable of teachers' “information technology self-efficacy” into the technology adoption model, and to verify its core role in the specific and highly situational field of art education, deepening the understanding of the complex psychological process of teachers' technology acceptance. The proposed “chain drive” mechanism provides a more refined explanatory framework for understanding the behavioral transformation from psychological initiation to deep application. At the practical level, the research conclusions provide direct and specific theoretical basis for designing truly “teacher-centered” AI assessment tools. The proposed design framework emphasizes that system development should not only pursue technological advancement, but should also strive to empower teachers and lower the adoption threshold through easy-to-use interaction, value explicitness, and stepped support, thereby effectively promoting the deep integration of intelligent technology into art teaching evaluation and helping to realize the paradigm transformation from experience-led to data-driven, and from result evaluation to process growth.

This study also has several limitations that need to be improved in future research. First of all, the sample size (N=103) is relatively limited, and most of the interviewees are “early adopters” who are familiar with AI technology. Although this provides a unique perspective for exploring the in-depth application mechanism, the conclusions need to be cautious when generalizing to a wider group of teachers with different technical foundations. Secondly, the study uses cross-sectional data. Although it can reveal the correlation between variables, it is difficult to strictly infer long-term causal effects. Finally, the system design framework proposed in this article is still at the conceptual level, and its actual effectiveness needs to be verified through prototype development, implementation and effect evaluation. Future research can go in the following directions: expand the sample size and diversity, conduct longitudinal tracking to capture the dynamic evolution of adoption behavior, develop a prototype system based on this framework, test, iterate and improve its effectiveness in practice through design research, and ultimately build a benign intelligent education ecosystem that can both empower teachers and continue to learn from teachers' practice.

References

[1] Yang, H. (2025). A new exploration of implementing "art smart education": Reflections on the application of artificial intelligence technology (AI technology) in primary school art teaching. *Jiazhang, 8*, 153–155.

[2] Lu, Q., & Cheng, J. (2018). An analysis of the impact of AI on art education and art teachers. *Journal of Puer University, 34*(1), 129–130.

[3] Zang, B. (2024). A preliminary study on AI-empowered teaching and reform of art theory courses in colleges and universities. *Shaanxi Education (Higher Education), 10*, 30–31. https://doi.org/10.16773/j.cnki.1002-2058.2024.10.009

[4] Jiang, Y., Fan, Y., & Liu, Z. (2025). Generative AI in Art Education: A Systematic Review of Research Trends, Tool Applications, and Outcomes (2019–2025). *Education Sciences*, *16*(1), 47.

[5] Peng, Y., Wang, X., Chen, J., Wang, Y., Dong, Z., Wang, Y., ... & Wang, W. (2025). The impact of AI-supported instructional interventions on students' creative thinking: a meta-analysis of 56 experimental studies. *Journal of Research on Technology in Education*, 1-25.

[6] Shen, T. (2024). Research on innovative development of integrated teaching of high school art and information technology under the background of "AI" empowerment. In Proceedings of the 11th Innovation Education Academic Conference — Technology and Education Development (pp. 20–22). Shenzhen Pengda Senior High School. https://doi.org/10.26914/c.cnkihy.2024.002395

[7] Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS quarterly*, *13*(3), 319-340.

[8] Albedah, F. (2025). Artificial Intelligence in Language Education: A Systematic Review of Multilingual Applications, Large Language Models, and Emerging Challenges. *Language Teaching Research Quarterly*, *49*, 247-268.

[9] Kim, K., & Kwon, K. (2024). Designing an inclusive artificial intelligence (AI) curriculum for elementary students to address gender differences with collaborative and tangible approaches. *Journal of Educational Computing Research*, *62*(7), 1617-1644.

[10] Jiang, Y., Fan, Y., & Liu, Z. (2025). Generative AI in Art Education: A Systematic Review of Research Trends, Tool Applications, and Outcomes (2019–2025). *Education Sciences*, *16*(1), 47.

[11] Luo, J., Zheng, C., Yin, J., & Teo, H. H. (2025). Design and assessment of AI-based learning tools in higher education: A systematic review. *International Journal of Educational Technology in Higher Education*, *22*(1), 42.

[12] Tsai, Y. S., & Gasevic, D. (2017, March). Learning analytics in higher education---challenges and policies: a review of eight learning analytics policies. In *Proceedings of the seventh international learning analytics & knowledge conference* (pp. 233-242).

[13] Mailizar, M., Almanthari, A., & Maulina, S. (2021). Examining teachers' behavioral intention to use E-learning in teaching of mathematics: An extended TAM model. *Contemporary educational technology*, *13*(2), ep298.

[14] Hazzan-Bishara, A., Kol, O., & Levy, S. (2025). The factors affecting teachers' adoption of AI technologies: A unified model of external and internal determinants. *Education and Information Technologies*, *30*(11), 15043-15069.

[15] Abulail, R. N., Badran, O. N., Shkoukani, M. A., & Omeish, F. (2025). Exploring the factors influencing AI adoption intentions in higher education: An integrated model of DOI, TOE, and TAM. *Computers*, *14*(6), 230.

[16] Sangwa, S., Mutabazi, P., & Muvunyi, J. B. (2025). AI-Enabled Framework for Program and Course Design in Higher Education.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.