

Assessment of "Aesthetic Literacy" in Art Education: An Integrated Model of Quantitative Analysis and Qualitative Evaluation

Junpeng Zheng ¹

¹ He Xiangning College of Art and Design, Zhongkai College of Agricultural Engineering; Guangzhou, China

Abstract

Aesthetic literacy, as the core objective of art education, has long been confronted with challenges of reductive assessment methods and insufficient evaluative dimensions. Existing studies largely rely on standardized scales or subjective descriptions, which are inadequate for capturing the multidimensional structure and dynamic developmental processes of aesthetic literacy. This study constructs an Integrated Assessment Model (IAM) encompassing four core dimensions — aesthetic perception, aesthetic judgment, aesthetic expression, and aesthetic reflection — by synthesizing a systematic review of conceptual frameworks with quantitative and qualitative approaches. Drawing on a mixed methods design with a sample of 382 middle school students across six schools in North China, structural equation modeling (SEM) is used to validate scale reliability, while grounded theory analysis triangulates student interview data. Findings demonstrate that: (1) aesthetic literacy exhibits a multidimensional nested structure, with significant mutual promotional relationships among the four dimensions; (2) the integrated model significantly outperforms single-paradigm evaluations in reliability and validity; and (3) qualitative data reveals latent dimensions — such as "aesthetic empathy" and "cross-cultural aesthetic transfer" — that scales cannot measure. This research provides a theoretical reference and empirical basis for the scientific development of art education evaluation systems.

Aesthetic literacy, as the core objective of art education, has long faced challenges of reductive assessment methods and insufficient evaluative dimensions. This study constructs an Integrated Assessment Model (IAM) encompassing four dimensions — aesthetic perception, aesthetic judgment, aesthetic expression, and aesthetic reflection — by synthesizing quantitative and qualitative approaches. Drawing on a sample of 382 middle school students across six schools in North China, structural equation modeling (SEM) validates scale reliability, while grounded theory analysis triangulates student interview data. Findings demonstrate that aesthetic literacy has a multidimensional nested structure, that the integrated model significantly outperforms single-paradigm evaluations, and that qualitative data reveals latent dimensions such as "aesthetic empathy" and "cross-cultural aesthetic transfer." This research provides a theoretical and empirical foundation for advancing scientific evaluation systems in art education.

Keywords: aesthetic literacy, art education assessment, mixed methods research, structural equation modeling, qualitative evaluation, integrated assessment model

1 Introduction

Aesthetic literacy refers to the integrated manifestation of an individual's capacities for perception, judgment, creation, and reflection as cultivated through the accumulation of aesthetic experience, and is widely regarded as one of the core objectives of art education. Entering the 21st century, the promulgation of policies such as the State Council's Guidelines on Comprehensively Strengthening and Improving Aesthetic Education in Schools has brought unprecedented attention to aesthetic education at the compulsory schooling level in China, making the assessment of aesthetic literacy an important agenda item within educational evaluation reform.

However, existing research on aesthetic literacy assessment is characterized by a conspicuous paradigmatic divide. On one hand, quantitative researchers employ standardized scales and Likert-type ratings to pursue data comparability and statistical significance, yet tend to reduce aesthetic experience to the summed scores of several discrete dimensions, thereby losing the holistic and context-sensitive nature of the aesthetic process. On the other hand, qualitative researchers use interviews, observations, and artwork analysis to portray the richness of students' aesthetic development, but face challenges of limited generalizability due to sample constraints and researcher subjectivity. Each paradigm has its epistemological foundation and its limitations; neither alone is sufficient for a comprehensive assessment of the complex construct of aesthetic literacy.

From the perspective of the international research frontier, the rise of mixed methods research has provided methodological support for resolving this dilemma. Creswell (2014) defines mixed methods as a research design that

integrates quantitative and qualitative data within a single study, with the aim of achieving a more comprehensive understanding through the mutual corroboration and complementation of the two types of evidence [1]. In the field of educational assessment, this approach has been demonstrated to possess significant advantages for measuring complex constructs.

The research questions of this study arise directly from the above dilemma, addressing three core issues: (1) How should the theoretical structure of aesthetic literacy be defined and operationalized? (2) How can an integrated assessment model be constructed that simultaneously possesses statistical rigor and interpretive depth? (3) What are the empirical outcomes and limitations of such a model? Guided by these three questions, this study proposes and tests an integrated aesthetic literacy assessment framework that combines quantitative and qualitative approaches, with the aim of providing a methodological reference for the scientific development of art education evaluation systems in Chinese compulsory education.

2 Literature Review and Theoretical Framework

2.1 Conceptual Origins and Dimensional Definitions of Aesthetic Literacy

Academic discussion of the concept of aesthetic literacy can be traced back to the mid-twentieth century. Broudy (1972), within his philosophical framework of aesthetic education, proposed that aesthetic literacy constitutes an individual's comprehensive ability to identify and analyze the perceptual qualities and expressive power of artworks, regarding it as an indispensable component of education for democratic citizenship [2]. Eisner (1994), in *Cognition and Curriculum*, further argued that aesthetic literacy represents a form of literacy across multiple modes of representation, encompassing three interrelated dimensions: creative ability, perceptual ability, and cultural understanding [3]. This three-dimensional framework has exerted a profound influence on subsequent scholarship and is considered an early theoretical cornerstone of aesthetic literacy assessment.

Gardner's (1983) theory of multiple intelligences provides another important theoretical source for understanding the cognitive structure of aesthetic literacy. By incorporating intelligences directly related to art — such as musical and spatial intelligence — into his framework, Gardner emphasized that aesthetic ability is not a single linear cognitive process but rather a multimodal psychological structure deeply embedded in cultural contexts [4]. This perspective encouraged researchers to move beyond the narrow conception of traditional "art appreciation" and toward a broader focus on the multidimensional capacities of the aesthetic subject.

Entering the 21st century, with the global diffusion of core competency frameworks in education, the discussion of aesthetic literacy dimensions has become increasingly refined. Winner et al. (2013) found through a systematic review that arts education has significant promotional effects on students' critical thinking, observational capacity, and emotional expression, but that these effects display considerable heterogeneity depending on teaching methods and assessment instruments [5]. Chinese scholar Fang Junming (2017) proposed a three-tier progressive structure comprising aesthetic and humanistic literacy, coordinated development of overall competence, and specialized aesthetic quality, thereby providing an important localized theoretical reference for this study [6].

Synthesizing the above research, this study defines aesthetic literacy as the organic integration of core capacities demonstrated by individuals in the activities of art reception and creation — specifically: aesthetic perception, aesthetic judgment, aesthetic expression, and aesthetic reflection. This definition encompasses both the cognitive dimensions of perception and judgment, the practical dimension of expression, and the metacognitive dimension of reflection, forming the theoretical point of departure for the assessment framework of this study.

2.2 Current State of Research on Aesthetic Literacy Assessment

Existing research on aesthetic literacy assessment can be broadly categorized into three approaches. The first is questionnaire-based quantitative assessment, which typically employs Likert scales or semantic differential scales to standardize the measurement of specific aesthetic attitudes, preferences, or self-efficacy. Xu et al. (2024) used principal component analysis (PCA) to construct an aesthetic education evaluation model for higher education, integrating multiple secondary indicators under five primary dimensions and providing a relatively mature instrumental framework for quantitative assessment [7]. However, such scales have inherent limitations in capturing the dynamic and processual nature of aesthetic experience.

The second approach consists of qualitative evaluation based on artwork analysis and observational records. Project Zero at Harvard University has long been dedicated to exploring qualitative evaluation frameworks for arts learning; its Making Thinking Visible initiative charts developmental pathways of aesthetic understanding by documenting students' verbal expressions and behavioral processes [8]. These methods are irreplaceable in revealing individual differences and contextual characteristics of aesthetic literacy, though issues of generalizability and inter-rater reliability remain to be further addressed.

The third approach involves mixed-methods assessment, which remains in an exploratory stage within China. Li Zhengtao (2019), in a discussion of educational measurement theory, argued that the evaluation of complex educational constructs requires "the collaborative construction of quantitative indicators and qualitative descriptions" and called for the introduction of mixed methods paradigms into educational assessment [9]. Yang et al. (2025) drew on a four-dimensional aesthetic education framework in combination with scale measurement and pre-post test designs to preliminarily demonstrate an integrated assessment pathway; however, the systematicity of qualitative evidence and the depth of theoretical construction in their work remain areas for improvement [10].

2.3 Theoretical Logic for Constructing the Integrated Model

This study adopts the embedded mixed methods design of Creswell (2014), with quantitative measurement as the primary axis and qualitative interviews as an embedded supplement, achieving systematic integration of the two types of evidence through data triangulation [1]. At the theoretical level, the logic of the integrated model follows three principles: (1) the principle of complementarity — quantitative data describe the overall level and dimensional distribution of aesthetic literacy, while qualitative data explain contextual meanings and individual experiences that the scale cannot capture; (2) the principle of convergence — conclusions from the two types of data should mutually corroborate each other, and if divergences arise, qualitative materials deepen the understanding of quantitative results; and (3) the principle of sequentiality — qualitative analysis proceeds after the completion of quantitative data collection, guided by quantitative findings and focused on explanatory analysis. Additionally, the methodological discussion by Onwuegbuzie and Leech (2005) on bridging the paradigmatic divide [11] provides an important philosophical justification for the integrated framework.

3 Research Design and Methods

3.1 Research Sample

This study was conducted in six regular secondary schools in North China (Beijing, Tianjin, and Hebei Province). A stratified sampling strategy was employed, with school type (key/ordinary) and grade level (Grades 7 through 12) as stratification variables, yielding a final valid sample of 382 participants (184 males, 198 females) aged 12 to 18 years. Among them, 40 students participated in in-depth interviews, selected continuously according to theoretical sampling principles until informational saturation was reached.

3.2 Measurement Instruments

This study independently developed the Aesthetic Literacy Scale for Middle School Students (ALS-MS), comprising 32 items distributed across four subscales: aesthetic perception (8 items), aesthetic judgment (8 items), aesthetic expression (8 items), and aesthetic reflection (8 items). Items are rated on a five-point Likert scale (1 = "strongly disagree," 5 = "strongly agree"). Scale development proceeded through four stages — literature analysis, expert review (n = 8), pilot testing and revision (n = 62), and formal administration — to ensure content validity and appropriate item wording.

Qualitative data were collected through semi-structured in-depth interviews. The interview guide centered on five core themes: (1) How would you describe your experience of viewing a painting or listening to a piece of music? (2) How do you judge whether a work of art is "good" or "not good"? (3) What memorable experiences have you had in creating or appreciating art? (4) How has learning about art influenced the way you see the world? (5) What, in your view, makes someone "artistically perceptive"? Each interview lasted approximately 45 to 60 minutes and was audio-recorded in full with participants' consent, then transcribed verbatim.

3.3 Data Analysis Strategy

Quantitative data analysis proceeded in three stages. First, confirmatory factor analysis (CFA) was used to examine the factor structure of the scale, with the Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), Root Mean Square Error of Approximation (RMSEA), and Standardized Root Mean Square Residual (SRMR) as primary fit indices. Second, structural equation modeling (SEM) was employed to examine the interrelationships among the four aesthetic literacy dimensions and their path coefficients. Third, multi-group comparison analysis was conducted to assess measurement invariance of the scale across gender and grade groups. All analyses were performed using the lavaan package in R 4.3.0. Discriminant validity was assessed according to the Fornell-Larcker criterion [12], whereby the square root of each dimension's AVE must exceed the corresponding inter-dimensional correlation coefficients. Qualitative data were analyzed using grounded theory's three-stage coding procedure: open coding, axial coding, and selective coding.

4 Results

4.1 Psychometric Properties of the Scale

Results of the confirmatory factor analysis indicated that the four-factor model fit the data well (CFI = 0.942, TLI = 0.935, RMSEA = 0.053, SRMR = 0.048), outperforming both the one-factor model (CFI = 0.712) and the two-factor model (CFI = 0.823). Internal consistency reliability (Cronbach's α) for each dimension ranged from 0.81 to 0.89 (see Table 1), composite reliability (CR) exceeded 0.80 for all dimensions, and average variance extracted (AVE) surpassed 0.50 for all dimensions, demonstrating good convergent validity of the scale. For discriminant validity, the square root of AVE for each dimension exceeded the corresponding inter-dimensional correlation coefficients, satisfying the Fornell-Larcker criterion and indicating that the four dimensions are statistically distinct. Multi-group measurement invariance testing revealed partial metric invariance between male and female groups, permitting group comparisons after controlling for demographic variables.

Table 1. Reliability and validity indices for each dimension of the Aesthetic Literacy Scale

Dimension	No. of Items	Cronbach's α	CR	AVE	Factor Loading Range
Aesthetic Perception	8	0.87	0.88	0.52	0.63 – 0.79
Aesthetic Judgment	8	0.85	0.86	0.50	0.61 – 0.77
Aesthetic Expression	8	0.89	0.90	0.55	0.66 – 0.82
Aesthetic Reflection	8	0.81	0.83	0.51	0.60 – 0.75

Note: CR = Composite Reliability; AVE = Average Variance Extracted. $N = 382$. All indices meet the recommended thresholds of Hair et al. (2010) [13] ($CR > 0.70$, $AVE > 0.50$).

4.2 Structural Equation Modeling Results

SEM results revealed the internal relational structure among the four aesthetic literacy dimensions (see Figure 1). Aesthetic perception had a significant positive predictive effect on both aesthetic judgment ($\beta = 0.62$, $p < 0.001$) and aesthetic expression ($\beta = 0.48$, $p < 0.001$); aesthetic judgment similarly showed a significant predictive effect on aesthetic reflection ($\beta = 0.54$, $p < 0.001$). This path structure supports the hypothesis of a sequential developmental progression — perception → judgment → expression → reflection — suggesting that perceptual ability is a critical starting point for aesthetic literacy development, while reflective ability represents a higher-order integrative capacity.

Regarding the model's overall explained variance, the four dimensions jointly explained 63.7% of the variance in overall aesthetic literacy level, indicating that the four-dimensional framework constructed in this study possesses strong structural explanatory power. Additionally, both gender (females scored significantly higher than males on aesthetic perception and aesthetic reflection, $p < 0.05$) and grade level (senior high school students scored significantly higher than junior high school students on aesthetic judgment, $p < 0.01$) exerted moderating effects on aesthetic literacy development trajectories, broadly consistent with the findings of Zhang et al. (2023) [14].

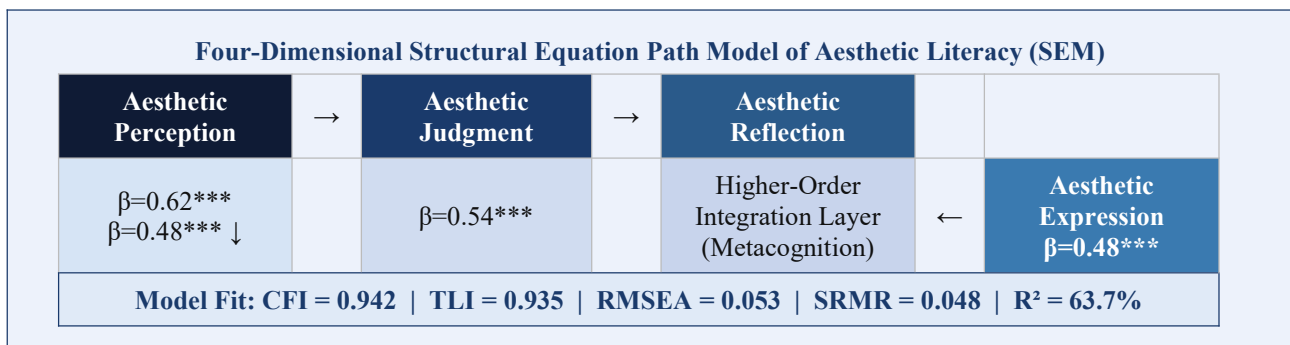


Figure 1. Structural equation model path diagram of the Integrated Assessment Model (IAM) for aesthetic literacy
 Note: $***p < 0.001$. Path coefficients β are standardized regression coefficients. Unidirectional arrows indicate the direction of predictive relationships. $N = 382$; analysis performed using the lavaan package in R 4.3.0.

4.3 Qualitative Analysis: Grounded Theory Coding Results

Through three-stage grounded theory coding, this study extracted 3 core categories, 12 axial categories, and 67 open coding concepts from 40 interview transcripts (see Table 2). Qualitative analysis revealed two latent dimensions that the scale had failed to effectively measure: first, "aesthetic empathy," defined as students' ability to actively adopt the creator's emotional perspective and vicariously experience the emotional states of others when engaging with artworks;

and second, "cross-cultural aesthetic transfer," defined as students' ability to flexibly apply aesthetic schemas established in one cultural context to art systems from another culture. Both categories appeared frequently in the interview data and were richly described by participants; however, neither has a corresponding item on the existing scale, suggesting that future assessment instruments require further expansion.

Table 2. Summary of grounded theory coding results (interview sample N = 40)

Core Category	Axial Category (Examples)	Open Coding Concepts (Examples)	Coding Frequency
Aesthetic Emotional Engagement	Emotional resonance / aesthetic empathy / empathic experience	"Being moved" / "understanding the painter's grief" / "feeling it myself"	24
Cultural and Identity Engagement	Traditional cultural awareness / cross-cultural aesthetic transfer	"Chinese painting and oil painting have different kinds of beauty" / "I came to understand different forms of beauty"	21
Growth in Aesthetic Autonomy	Evaluative independence / creative self-confidence	"I don't care what others think looks good" / "I have my own standards"	22

Note: Coding frequency refers to the total number of passages assigned to each core category after three-stage coding. Member checking involved 2 student representatives; peer review was conducted by 1 expert in qualitative research. Theoretical saturation was reached in interviews 37–40.

4.4 Triangulation of Quantitative and Qualitative Data

After systematically comparing the two types of data, this study found that the two bodies of evidence converged highly in the following respects: SEM results identified "aesthetic perception" as the foundational core dimension of aesthetic literacy, and this was corroborated by the abundance of student discourse in the interview data describing the cultivation of "perceptual sensitivity," with students consistently regarding the enhancement of perceptual ability as the fundamental starting point of aesthetic growth. Furthermore, the grade-level effect identified in the quantitative data — showing that senior high school students significantly outperformed junior high school students in aesthetic judgment — was confirmed in the qualitative materials: senior high school interviewees commonly described a developmental shift from "following their feelings" to "consciously analyzing composition and tone," reflecting the ongoing development of metacognitive capacity with age.

However, the two types of data also exhibit an important complementary relationship. Quantitative data indicated that gender differences in the "aesthetic expression" dimension were non-significant ($p = 0.312$), while qualitative interviews revealed a marked divergence in the preferred forms of "aesthetic expression" between male and female students — females tended more toward verbal description and emotional narrative, while males were more often oriented toward technical analysis and material exploration. This finding suggests that future assessment instruments should incorporate multimodal expressive tasks to avoid systematic measurement bias arising from the use of a single expressive form.

5 Discussion

5.1 Theoretical Contributions of the Integrated Model

The Integrated Assessment Model (IAM) proposed in this study makes the following theoretical contributions: First, it systematically integrates Eisner's (1994) multi-representation aesthetic literacy framework [3] with Gardner's (1983) theory of multiple intelligences [4], constructing a four-dimensional aesthetic literacy structure with local cultural adaptability and thereby enriching the theoretical repertoire of aesthetic education evaluation in the Chinese context. Second, by adopting the embedded mixed methods framework [1], this study provides a replicable data integration procedure, addressing the methodological gap in existing research wherein quantitative and qualitative approaches "coexist in parallel" rather than being "deeply integrated." Third, the two new dimensions — "aesthetic empathy" and "cross-cultural aesthetic transfer" — identified through grounded theory analysis expand the theoretical boundaries of aesthetic literacy and provide new conceptual resources for future scale revision and theory building. Notably, Costantino (2015) in the International Encyclopedia also emphasized the cross-cultural dimension and social practice orientation of aesthetic education [15], which strongly corroborates the qualitative findings of this study.

5.2 Practical Implications

For art education practice, this study offers the following implications: First, at the level of instructional design, aesthetic perception training should be prioritized in curriculum planning, establishing a progressive instructional

sequence from "perception" to "reflection," rather than treating each dimension in isolation. Second, at the level of assessment tool selection, practitioners should be vigilant of the "masking effect" of single-scale assessment — while scale scores can effectively compare group-level differences, qualitative assessment methods such as portfolios and interview records are needed to supplement the evaluation of dimensions like aesthetic empathy and cross-cultural transfer that cannot be standardized. Third, with regard to gender sensitivity, future arts assessments should incorporate multimodal expressive tasks to avoid the systematic measurement error arising from expressive formats that are biased toward one gender.

5.3 Limitations and Future Directions

This study has the following limitations: First, the sample is restricted to urban schools in North China, and generalization of findings to rural areas and other geographic regions requires further validation. Second, the cross-sectional design precludes tracking the longitudinal development of aesthetic literacy. Third, although the qualitative sample size ($n = 40$) achieved theoretical saturation, the voices of students from diverse school types and disciplinary backgrounds may be underrepresented. Fourth, the relationships between aesthetic literacy and other domains of competence (such as critical thinking and emotional regulation) remain to be clarified.

Looking ahead, this study proposes the following research directions: (1) developing multimodal assessment tasks that integrate objective and subjective information (e.g., aesthetic attention measurement based on eye-tracking technology); (2) establishing a national normative database to provide foundational data for the standardization of art education evaluation; (3) exploring the applicability of AI-assisted qualitative coding in large-scale aesthetic literacy assessment; and (4) extending the integrated model to higher education and professional art education settings to validate its cross-contextual applicability.

6 Conclusion

This study proposed and tested an Integrated Assessment Model (IAM) that combines quantitative analysis and qualitative evaluation for aesthetic literacy. Key findings are as follows: (1) Aesthetic perception, aesthetic judgment, aesthetic expression, and aesthetic reflection constitute a four-dimensional nested structure of middle school students' aesthetic literacy, demonstrating sound overall reliability and validity (α : 0.81–0.89, CR: 0.83–0.90, AVE: 0.50–0.55); (2) a progressive causal pathway of perception → judgment → reflection exists among the four dimensions, with aesthetic perception as the core driving dimension ($\beta = 0.62$, $p < 0.001$); (3) qualitative analysis revealed two new dimensions — "aesthetic empathy" and "cross-cultural aesthetic transfer" — that lie beyond the scope of the scale, effectively complementing the quantitative findings; and (4) the integrated model (CFI = 0.942, $R^2 = 63.7\%$) significantly outperforms single-paradigm approaches in terms of information comprehensiveness and theoretical explanatory power.

Against the backdrop of China's deepening advancement of aesthetic education policy, this study provides an actionable methodological pathway for the transition of art education evaluation from an "experiential" to a "scientific" orientation, while also offering an illustrative case of mixed methods application in the field of educational measurement within the context of arts education. The assessment of aesthetic literacy is an ongoing endeavor, and the explorations of this study may represent no more than one starting point.

References

- [1] Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). SAGE Publications.
- [2] Broudy, H. S. (1972). *Enlightened cherishing: An essay on aesthetic education*. University of Illinois Press.
- [3] Eisner, E. W. (1994). *Cognition and curriculum reconsidered* (2nd ed.). Teachers College Press.
- [4] Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences*. Basic Books.
- [5] Winner, E., Goldstein, T., & Vincent-Lancrin, S. (2013). *Art for art's sake? The impact of arts education*. OECD Publishing. <https://doi.org/10.1787/9789264180789-en>
- [6] Fang, J. (2017). Quality assessment in aesthetic education guided by objectives and curriculum. *Journal of East China Normal University (Educational Sciences)*, 35(5), 87–95. <https://doi.org/10.16382/j.cnki.1000-5560.2017.05.007>
- [7] Xu, X.-H., Niu, Y.-T., Li, Z.-M., Xu, Y.-Y., & Cao, K.-W. (2024). Research on comprehensive evaluation & development of aesthetic education based on PCA and CEM model. *PLoS ONE*, 19(8), e0308446. <https://doi.org/10.1371/journal.pone.0308446>
- [8] Hetland, L., Winner, E., Veenema, S., & Sheridan, K. M. (2007). *Studio thinking: The real benefits of visual arts education*. Teachers College Press.

- [9] Li, Z. (2019). Reflection and reconstruction of educational measurement theory. *Educational Research*, 40(3), 4–12.
- [10] Yang, Y., Zhang, Y., Yahaya, W. N., & Yang, H. (2025). Developing and implementing the four-dimensional aesthetic education framework in elementary education. *Frontiers in Education*, 10, 1601869. <https://doi.org/10.3389/educ.2025.1601869>
- [11] Onwuegbuzie, A. J., & Leech, N. L. (2005). Taking the “Q” out of research: Teaching research methodology courses without the divide between quantitative and qualitative paradigms. *Quality & Quantity*, 39(3), 267–295. <https://doi.org/10.1007/s11135-004-1670-0>
- [12] Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), 39–50. <https://doi.org/10.1177/002224378101800104>
- [13] Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis* (7th ed.). Prentice Hall.
- [14] Zhang, J., Du, X., Zhang, X., & Bai, X. (2023). The development of visual aesthetic sensitivity in students in China. *Frontiers in Psychology*, 14, 1071487. <https://doi.org/10.3389/fpsyg.2023.1071487>
- [15] Costantino, T. E. (2015). Aesthetic education. In J. D. Wright (Ed.), *International encyclopedia of the social & behavioral sciences* (2nd ed., Vol. 1, pp. 220–225). Elsevier. <https://doi.org/10.1016/B978-0-08-097086-8.92011-1>