Development of a Psychometric Instrument for Creativity Potential and Art Appreciation

E.W Suprihatin Dyah Pratamawati[1], Sri Rahayuningsih[2*], Farah Damia Mohd Nasir[3]

[1*]Program Studi Seni dan Desain, Fakultas Sastra, Universitas Negeri Malang, Malang, Indonesia
[2]Program Studi Pendidikan Dasar, Sekolah Pascasarjana, Universitas Negeri Malang
[3]Faculty of Educational Studies, Universiti Tun Abdul Razak, Petaling Jaya, Selangor, Malaysia

Article Info

Article history:
Received Aug 22, 2025
Accepted Sep 22, 2025
Published Online Oct 23, 2025

Keywords:
Creativity potential
Art appreciation
Instrument development
Psychometric analysis
Arts education

ABSTRACT

This study aims to develop a valid and reliable instrument to measure creativity potential and art appreciation among students enrolled in the Arts and Culture Teacher Professional Education Program (PPG). The research was driven by the limited availability of assessment tools that comprehensively measure non-technical competencies, particularly creative thinking and aesthetic responsiveness, which are essential components in arts education. Employing a Research and Development (R&D) approach, this study adopted a modified Borg and Gall model consisting of seven stages, including needs analysis, blueprint formulation, item development, expert judgment for content validation, pilot testing, revisions, and large-scale field testing. Data were collected using a Likert-scale questionnaire and performance-based assessment tasks. Content validity was analyzed using Aiken's V, while construct validity was tested through Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA). Reliability was established using Cronbach's Alpha, McDonald's Omega, and Intraclass Correlation Coefficient (ICC). The findings confirmed that the developed instrument met strong psychometric standards, with factor loadings aligned with theories of creativity and contemporary aesthetics. The instrument effectively captures five dimensions of creativity (fluency, flexibility, originality, elaboration, and creative self-efficacy) and five dimensions of art appreciation (formal perception, interpretation, evaluation, transformational reflection, and tolerance of ambiguity). This instrument can serve as a diagnostic and evaluative tool and provide pedagogical insights for enhancing creative and reflective learning in arts education.

Corresponding Author:
Sri Rahayuningsih,
Program Studi Pendidikan Dasar,
Sekolah Pascasarjana,
Universitas Negeri Malang, Malang, Indonesia,
Jalan Semarang No. 5, Kota Malang, Malang, Indonesia
Email: *srirahayuningsih.pasca@um.ac.id*

How to cite: Pratamawati, E. S. D., Rahayuningsih, S., & Nasir, F. D. M. (2025). Development of a Psychometric Instrument for Creativity Potential and Art Appreciation. Jurnal Riset Dan Inovasi Pembelajaran, 5(3), 1017–1028. https://doi.org/10.51574/jrip.v5i3.3899

Development of a Psychometric Instrument for Creativity Potential and Art Appreciation

1. Introduction

The development of arts education in higher education requires students not only to master technical skills but also to acquire non-technical competencies that are essential in the 21st century, such as creativity and art appreciation. Creativity is recognized as a central competency for adapting to rapid change, generating innovation, and solving complex problems (Amabile & Pratt, 2016). Theoretically, creativity emerges from the interaction among domain-relevant knowledge, creative thinking skills, motivation, and supportive socio-environmental factors, as explained in the Dynamic/Componential Model. Recent developments in creativity research emphasize cognitive mechanisms, suggesting that the core of creativity lies in the ability to navigate semantic networks and activate remote associations to generate original ideas (Beaty & Kenett, 2023). Thus, measuring students' creative potential is not only essential for evaluating learning outcomes but also necessary for designing learning strategies that stimulate divergent and innovative thinking.

In addition to creativity, art appreciation is an equally important non-technical competence in arts education. It enables students to perceive, interpret, and evaluate works of art in meaningful ways. Contemporary theoretical frameworks such as the Vienna Integrated Model of Art Perception (VIMAP) highlight that aesthetic experience results from the integration of perceptual processes (bottom-up) with cognitive, emotional, and contextual influences (top-down) (Pelowski et al., 2017). This is reinforced by aesthetic cognitivism, which views art appreciation as a cognitive activity intertwined with emotional engagement and reflective interpretation (Christensen et al., 2023). Furthermore, aesthetic experiences hold transformative potential and may trigger cognitive discrepancy, insight, and self-reflection (Pizzolante et al., 2024), evoke feelings of awe and meaning-making (Sawada et al., 2024), and contribute to personal well-being through emotional elevation (Mattheiss & Herrmann, 2024).

However, despite the widely acknowledged importance of creativity and art appreciation in arts education, current assessment practices at the university level remain dominated by instruments focused solely on cognitive achievement or technical skill mastery. Instruments designed specifically to measure creativity and art appreciation are still limited in both availability and psychometric quality. Previous measures such as the Creative Self-Efficacy Scale (Karwowski et al., 2024) and aesthetic preference scales (Christensen et al., 2023) tend to assess only a single construct and fail to capture the multidimensional nature of creativity or the depth of art appreciation.

In this regard, several methodological gaps still persist. First, existing instruments tend to be partial in nature, focusing only on one aspect of creativity or art appreciation without integrating both constructs into a single comprehensive framework. Second, most existing instruments rely primarily on self-report scales and rarely combine perceptual, cognitive, and performative assessment techniques (Li et al., 2024). Third, emotional and transformative dimensions such as awe, meaning-making, and reflective transformation are widely discussed in art appreciation literature (Sawada et al., 2024; Pizzolante et al., 2024), yet remain rarely operationalized. Finally, most instruments are not contextualized to the domain-specific characteristics of arts and culture students in teacher education programs.

Preliminary diagnostic findings from students enrolled in the Arts and Culture Teacher Professional Education (PPG) Program reveal low levels of creative potential and limited depth of art appreciation. Most students generate only a limited number of ideas, demonstrate low confidence in presenting original work, and tend to remain at a superficial descriptive level when interpreting artworks. They often struggle to construct deeper meanings or provide

justified evaluative reasoning. This situation indicates a gap between expected learning outcomes and actual student performance. Therefore, there is a strong need for a valid, reliable, and contextually relevant instrument capable of measuring both creativity and art appreciation comprehensively.

Problem Statements

1. How can a valid and reliable instrument be developed to measure the creativity potential of students in the Arts and Culture Study Program?
2. How can a valid and reliable instrument be developed to measure the art appreciation potential of students in the Arts and Culture Study Program?
3. How do the results of validity and reliability testing demonstrate the instrument's ability to comprehensively describe students' creativity potential and art appreciation?

Research Objectives

1. To develop an instrument that can measure the creativity potential of students in the Arts and Culture Study Program with high validity and reliability.
2. To develop an instrument that can measure the art appreciation potential of students in the Arts and Culture Study Program accurately and consistently.
3. To produce an instrument that provides a comprehensive description of students' creativity potential and art appreciation, which can serve as a reference for lecturers and educational institutions in designing more effective arts learning strategies.

2. Method

Research Design This study employed a Research and Development (R&D) approach using a modified version of the Borg and Gall model. This model was selected because it provides a systematic procedure for developing educational products, including assessment instruments. From the ten original steps proposed by Borg and Gall, this study adapted seven essential stages, namely: (1) needs analysis, (2) instrument blueprint development, (3) item construction, (4) content validity through expert judgment, (5) small-scale try-out, (6) instrument revision, and (7) large-scale field testing to examine construct validity and reliability. These stages were simplified to align with the context of psychometric instrument development in arts education.

This approach is consistent with instrument development practices reported in international publications. For example, the development of the Creative Self-Efficacy (CSE) Scale followed procedures including content validation, pilot testing, and construct validation using Confirmatory Factor Analysis (CFA), which resulted in a stable factor structure and high reliability (Karwowski et al., 2024). Similarly, the Trait Creative Potential Questionnaire (TCPQ-12) was developed through literature review, indicator formulation, and validity and reliability testing, making it suitable for measuring creative potential (Primi et al., 2021).

In the field of art appreciation, corresponding instrument development also follows similar patterns. The Aesthetic Processing Preference Scale (APPS) was developed through expert validation, pilot testing, and both Exploratory and Confirmatory Factor Analyses, resulting in a three-dimensional structure related to individuals' aesthetic processing preferences (Muth et al., 2023). Likewise, the Art Criticism Assessment Rubric (ACAR) was constructed to assess students' responses to artworks using systematic expert validation, pilot testing, and revisions, enabling comprehensive evaluation of art appreciation skills (Kang & Kim, 2017).

Additionally, studies related to creativity and aesthetics assessment in specific domains offer relevant methodological references. The Scale of Aesthetics and Creativity in Chess (SACC) applied the Many-Facet Rasch Model to examine unidimensionality, validity, and reliability (García-Madruga et al., 2025), while the Aesthetics of Everyday Life Scale (AELSA)

was developed through systematic stages, including item generation, expert validation, and reliability testing across different age groups (Izadi-Avanji et al., 2024). In higher education, the Postgraduate Research Innovation Ability Scale demonstrated appropriate construct validity (EFA and CFA) and high reliability, serving as a useful reference for measuring non-technical competencies in students (Li et al., 2023).

Referring to these studies, the modified Borg and Gall model used in this research emphasizes psychometric validation, including content validity using Aiken's V, construct validity through EFA and CFA, and internal reliability using Cronbach's Alpha. This strategy ensures that the developed instrument is not only valid in terms of content but also demonstrates strong internal consistency, making it feasible for broad application in measuring creativity and art appreciation in higher education.

Research Subjects

The subjects involved in this instrument development consisted of three groups selected purposively according to the needs of each development stage: experts (expert judgment), small-scale try-out participants, and large-scale field test participants. At the content validation stage, two senior lecturers with expertise in arts education, assessment, and evaluation were involved. These experts were selected based on criteria such as more than ten years of teaching experience, publications in arts education and assessment, and active involvement in academic activities related to assessment. Their role was to evaluate item relevance, clarity, and alignment with the constructs of creativity and art appreciation, making expert validation a crucial stage to ensure adequate content validity prior to field implementation.

The next stage involved 30 students from the Arts and Culture Teacher Professional Education (PPG) Program. Selected through purposive sampling, the small-scale try-out aimed to evaluate item readability, language clarity, item difficulty level, and time allocation for completion. The results from this stage were used to revise items considered ambiguous or unrepresentative. Following revisions, a large-scale field test was conducted with 150 students from the same program at Universitas Negeri Malang. Data collected at this stage were used to analyze construct validity through Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA), as well as to examine internal reliability using Cronbach's Alpha. The number of respondents met the minimum requirement for factor analysis five times the number of instrument items (Hair et al., 2019) ensuring statistical rigor. With this subject design, the study aimed to produce an instrument that is valid in content (expert judgment), practical (pilot testing), and psychometrically robust (large-scale testing).

Research Procedure

The first stage was needs analysis conducted through literature review and preliminary diagnostics. The literature review focused on creativity theories such as the Dynamic Componential Model (Amabile & Pratt, 2016), Associative Thinking Account (Beaty & Kenett, 2023), and Creative Self-Efficacy (Abulela et al., 2024; Liu et al., 2025), along with theories of art appreciation such as the VIMAP Model (Pelowski et al., 2017), Aesthetic Cognitivism (Christensen et al., 2023), and Transformative Aesthetic Experience (Pizzolante et al., 2024). Preliminary diagnostics showed that students had low levels of creativity, particularly in fluency, flexibility, and creative self-efficacy, as well as weaknesses in meaning-making, reflection, and tolerance for ambiguity in art appreciation reinforcing the need for a more comprehensive instrument. The second stage involved developing the instrument blueprint based on theoretical frameworks and empirical findings. Creativity was conceptualized into five dimensions: fluency, flexibility, originality, elaboration, and creative self-efficacy. Art appreciation consisted of five dimensions: formal perception, interpretation, evaluation, transformative reflection, and tolerance of ambiguity. These dimensions were then operationalized into specific indicators and item plans.

The third stage was item development using a five-point Likert scale and performance-based tasks to assess divergent thinking (fluency, flexibility, originality, elaboration) and art appreciation (interpretation, evaluation, reflection). The items were designed to reflect both theoretical constructs and empirical characteristics of art education students. The fourth stage was expert judgment conducted by three experts in arts education and assessment, who evaluated item clarity, relevance, and appropriateness using a four-point rating scale. Aiken's V index was used to analyze content validity and identify items requiring revision. The fifth stage was a small-scale try-out involving 30 students to evaluate clarity and practicality. Ambiguous items were revised based on feedback. The sixth stage involved instrument revision based on expert and participant responses to ensure clarity, relevance, and comprehensivenessespecially in dimensions previously underrepresented, such as creative self-efficacy and transformative reflection. The final stage was the large-scale try-out involving 150 students to test construct validity using EFA and CFA and internal reliability using Cronbach's Alpha and McDonald's Omega. The instrument produced at this stage was expected to be valid, reliable, and ready for implementation in arts education contexts.

Data Collection Techniques

Data were collected through several techniques aligned with the stages of instrument development to ensure validity, reliability, and comprehensiveness: (1) Literature Review and Documentation, conducted at the early stage to establish the theoretical foundation of the instrument using Scopus-indexed journal articles, academic books, and arts education curriculum documents, focusing on creativity and art appreciation theories; (2) Preliminary diagnostic interviews; (3) Semi-structured interviews with lecturers and students explored empirical needs and current conditions related to creativity and art appreciation in the PPG program; (3) Likert-scale questionnaire, the main instrument consisted of 30 items (15 creativity and 15 art appreciation) administered in both small and large-scale try-outs; (4) Performance tasks, students completed art interpretation or evaluation tasks. Responses were scored using performance rubrics covering fluency, flexibility, originality, elaboration, interpretation, evaluation, and reflection; (5) Expert validation sheets, expert judgment data were collected using validation sheets rated on a 4-point scale, analyzed using Aiken's V.

Data Analysis Techniques

Data analysis followed psychometric procedures:

1. **Content Validity Analysis**

   Aiken's V was used to calculate content validity based on expert ratings:

$$V = \frac{\sum(ri - lo)}{\left(n(c-1)\right)}$$

   Criteria:

   $V \geq 0.80$ = very good (accepted)

   $0.70 \leq V < 0.80$ = acceptable with revision

   $V < 0.70$ = requires major revision or elimination

2. Small-Scale Try-Out Analysis

Descriptive analysis was conducted to evaluate item readability, clarity, and response patterns.

3. Construct Validity Analysis

   Exploratory Factor Analysis (EFA): using KMO and Bartlett tests, factor extraction via Principal Axis Factoring, rotation with Varimax/Oblimin, factor loading ≥ 0.40. Confirmatory Factor Analysis (CFA): using SEM with model fit criteria: $\chi^2/df \leq 3$, RMSEA ≤ 0.08, CFI ≥ 0.90, TLI ≥ 0.90, SRMR ≤ 0.08.

4. Reliability Analysis

   Cronbach's Alpha ($\alpha \geq 0.70$)

   McDonald's Omega ($\omega \geq 0.70$)

5. Performance Task Analysis

   Scored with rubrics and analyzed using descriptive statistics and Inter-rater reliability via Intraclass Correlation Coefficient (ICC).

3. Research Findings

The results of interviews with lecturers and students revealed that existing assessment instruments in arts education primarily measure technical skills, while essential non-technical competencies such as creativity (fluency, flexibility, originality, elaboration, and creative self-efficacy) and art appreciation (formal perception, interpretation of meaning, evaluative judgment, transformative reflection, and tolerance of ambiguity) have not been adequately addressed. Preliminary diagnostic data from 40 students showed low levels of creative potential, with an average fluency score of only two ideas per artwork and creative self-efficacy mean score of 2.8 on a 1-5 scale. Similarly, art appreciation scores were low on reflection (M = 2.7) and tolerance for ambiguity (M = 2.9). These findings emphasize the urgency of developing a more comprehensive instrument capable of measuring these constructs effectively.

Following the needs analysis stage, the next step in developing the instrument was to examine its content validity to ensure that each item accurately represented the intended construct. Content validity is a crucial requirement in the development of measurement instruments, as it provides initial evidence that the items are relevant, clear, and theoretically aligned with the domains being measured. In this study, content validity was established through expert judgment conducted by specialists in arts education and educational assessment. Each expert evaluated the relevance, clarity, and representativeness of the items using a four-point rating scale. The results of expert evaluations were then analyzed using Aiken's V coefficient to quantify the degree of agreement among experts regarding the suitability of each item. The summary of Aiken's V analysis is presented in Table 1 below.

Table 1. Summary of Aiken's V Coefficient by Dimension

Dimension	Number of Items	Range of V	Mean V	Criteria
Creativity	15	0.75–0.95	0.86	Valid

Dimension	Number of Items	Range of V	Mean V	Criteria
Art Appreciation	15	0.72–0.93	0.84	Valid
Total Instrument	30	0.72–0.95	0.85	Valid

Most items were rated highly valid by experts (Aiken's V ≥ 0.80). A few items with V values between 0.70 and 0.79 were revised for clarity, such as items related to "providing alternative interpretations" and "engaging with ambiguous artworks." No items scored below 0.70; therefore, all items were retained.

After establishing content validity, the next stage was to evaluate the practicality and readability of the instrument through a small-scale try-out. This stage aimed to identify potential issues related to item clarity, linguistic ambiguity, and response consistency before administering the instrument to a larger sample. The small-scale try-out involved 30 students who were representative of the target population, and their feedback was used to refine the instrument. Several items required editorial adjustments to improve their contextual relevance and clarity. Examples of item revisions are presented in Table 2.

Table 2. Sample Item Revisions from Small-Scale Try-Out (N = 30)

No.	Item Before Revision	Issue Identified	Revised Item
K2	"I can mention more than three interpretations of a painting."	Difficult to understand, overly specific	"I can provide several different interpretations of an artwork."
A13	"I enjoy artworks that are difficult to understand."	Too brief and ambiguous	"I remain engaged even when an artwork is difficult to understand."

The small-scale try-out indicated that most items were understandable for students. However, several items required revision to improve contextual clarity. Revisions were made based on input from students and expert validators.

Following revisions from the pilot testing, the revised instrument was administered to a larger group of respondents to examine its psychometric properties more rigorously. The large-scale try-out aimed to evaluate construct validity and reliability to determine whether the instrument could be used as a valid measurement tool in educational research. Construct validity was assessed through both Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA), while internal consistency reliability was tested using Cronbach's Alpha and McDonald's Omega coefficients. The results of construct validity testing are presented in the following subsections.

Table 3. EFA Results (N = 150)

Statistical Indicator	Value	Criteria	Conclusion
Kaiser-Meyer-Olkin (KMO)	0.89	≥ 0.80	Adequate
Bartlett's Test of Sphericity	$\chi^2 = 1530$, $p < .001$	Significant	Suitable for EFA
Factor Loading	0.42–0.81	≥ 0.40	Valid

The instrument demonstrated a factor structure consistent with theoretical expectations, with five factors for creativity and five for art appreciation. All items met the minimum factor loading requirement (≥ 0.40). To further confirm the factor structure obtained from the EFA, Confirmatory Factor Analysis (CFA) was conducted. CFA was used to assess whether the measurement model demonstrated an acceptable fit with the empirical data and whether the observed variables corresponded to the hypothesized latent constructs. Several model fit indices were examined, including Chi-square/df, RMSEA, CFI, TLI, and SRMR. The results of the CFA are presented in Table 4.

Table 4. CFA Results (N = 150)

Fit Index	Value	Criterion	Interpretation
χ^2/df	2.11	≤ 3	Good fit
RMSEA	0.056	≤ 0.08	Good fit
CFI	0.93	≥ 0.90	Good fit
TLI	0.91	≥ 0.90	Good fit
SRMR	0.047	≤ 0.08	Good fit

The measurement model demonstrated a good fit with the empirical data, confirming that creativity and art appreciation are valid multidimensional constructs. After confirming the construct validity of the instrument, the next step was to evaluate its reliability to ensure measurement consistency. Internal consistency reliability was assessed using Cronbach's Alpha (α) and McDonald's Omega (ω), both of which are widely used indicators in educational and psychological measurement. Reliability testing was conducted for each dimension of the instrument. The results are summarized in Table 5.

Table 5. Reliability Results

Dimension	Cronbach's Alpha (α)	McDonald's Omega (ω)	Interpretation
Creativity	0.87	0.88	Reliable
Art Appreciation	0.85	0.86	Reliable

All instrument dimensions showed reliability coefficients above 0.85, indicating excellent internal consistency. In addition to the self-report questionnaire, this study also employed performance-based tasks to provide a more comprehensive assessment of creativity and art appreciation. Students' responses to performance tasks were evaluated using an analytic rubric, and inter-rater reliability was calculated using the Intraclass Correlation Coefficient (ICC) to ensure scoring consistency among raters. The results of inter-rater reliability analysis are presented in Table 6.

Table 6. Results of Inter-Rater Reliability Analysis

Dimension	ICC	Interpretation
Fluency	0.80	Very Good
Flexibility	0.82	Very Good
Originality	0.79	Good
Elaboration	0.81	Very Good
Interpretation	0.83	Very Good
Evaluation	0.85	Very Good
Reflection	0.82	Very Good

The performance rubric demonstrated high inter-rater reliability (ICC > 0.79), indicating consistent scoring between assessors. Overall, the findings indicate that the developed instrument demonstrates strong psychometric properties. However, further analysis of student performance data reveals important implications for arts education, particularly related to enhancing creative self-efficacy and deeper levels of artistic interpretation. These findings are discussed in the following section in relation to relevant theories and previous research.

4. Discussion

The inter-rater reliability analysis for the performance-based assessment showed values ranging from good to very good, indicating that the rubric used in this study is suitable for objectively evaluating students' abilities. Overall, the findings demonstrate that the developed instrument is capable of measuring students' creativity potential and art appreciation with adequate content validity, construct validity, and reliability. However, results from the small-

scale try-out and preliminary diagnostic phase revealed that students still exhibited weaknesses in several key dimensions of creativity, particularly fluency, flexibility, and creative self-efficacy. In terms of art appreciation, the lowest scores appeared in the dimensions of meaning interpretation, self-reflection, and tolerance of ambiguity.

These findings align with previous research showing that creativity is not solely related to the ability to generate new ideas but is also strongly influenced by one's confidence in their creative capabilities, known as creative self-efficacy (Tao et al., 2024). Similarly, Capron Puozzo and Audrin (2021) found that although pedagogical interventions do not always lead to significant increases in creative self-efficacy, they do foster students' understanding of the creative process and enhance awareness of the importance of creativity in learning. This suggests that creative self-efficacy should be intentionally developed through reflective and task-based learning strategies.

In relation to art appreciation, students demonstrated stronger abilities in formal perception and technical evaluation but showed weaknesses in deeper meaning-making and transformative reflection. According to the Vienna Integrated Model of Art Perception (VIMAP), meaningful art appreciation requires the integration of bottom-up perceptual processes with top-down cognitive and emotional engagement (Pelowski et al., 2017). Supporting this view, Iosifyan (2021) reported that theory of mind the ability to understand the intentions and emotions behind an artwork plays a crucial role in enhancing aesthetic appreciation. These findings suggest that students' limitations in interpretive and reflective abilities may be due to a lack of learning experiences that promote deeper engagement with artworks in social, cultural, and emotional contexts.

The high construct validity of the instrument, confirmed through both Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA), indicates that the factor structure aligns with contemporary theories of creativity and art appreciation (Amabile & Pratt, 2016; Beaty & Kenett, 2023). Supported by strong KMO values, acceptable factor loadings, and good fit indices, the instrument successfully measures creativity and art appreciation as multidimensional constructs. Furthermore, the high reliability coefficients (Cronbach's Alpha and McDonald's Omega > 0.85) and strong inter-rater reliability in performance analysis (ICC > 0.80) demonstrate the robustness of the instrument. These results reinforce the importance of combining self-report measures with performance-based assessments to obtain a more comprehensive evaluation of students' creative and aesthetic competencies (Capron Puozzo & Audrin, 2021; Tao et al., 2024).

In conclusion, this study confirms that the developed instrument meets strong psychometric standards and can be considered a valid and reliable tool for assessing creativity and art appreciation among students in higher education. Nonetheless, the findings also highlight important pedagogical implications, particularly the need to design learning strategies that foster creative confidence, reflective thinking, and tolerance for ambiguity. Such instructional efforts could include interpretive tasks involving complex artworks or collaborative activities that emphasize meaning negotiation and critical dialogue. Therefore, beyond its role as an assessment tool, this instrument holds potential as a reference for designing more effective educational interventions that support the holistic development of creativity and art appreciation.

5. Conclusion

This study successfully developed an instrument for measuring creativity potential and art appreciation among students in the Arts and Culture Study Program, demonstrating strong psychometric qualities. The instrument, consisting of 30 Likert-scale items complemented by performance-based tasks, was proven to have high content validity, as indicated by Aiken's V

coefficients which were predominantly within the very good category. The instrument also demonstrated construct validity, supported by the results of Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA), which confirmed the multidimensional structure of creativity (fluency, flexibility, originality, elaboration, and creative self-efficacy) and art appreciation (perception, interpretation, evaluation, reflection, and tolerance of ambiguity). Furthermore, the instrument showed high reliability, evidenced by Cronbach's Alpha, McDonald's Omega, and strong inter-rater reliability based on ICC values from the performance rubric.

The findings also revealed that students showed weaknesses in certain domains, particularly creative self-efficacy and the reflective and ambiguity tolerance dimensions of art appreciation. These results highlight the importance of developing instructional strategies that emphasize interpretive, reflective, and collaborative learning experiences. Such approaches can help students move beyond surface-level analysis and foster deeper engagement with artworks, enabling them to interpret meaning, conduct self-reflection, and appreciate diverse perspectives. Therefore, the developed instrument is not only valuable as an academic assessment tool but also serves as a practical reference for lecturers and educational institutions in designing pedagogical interventions that holistically support the development of creativity and art appreciation in higher education.

6. Author Contributions

E.W.S.D.P. conceptualized the study, designed the research methodology, and led the development of the instrument and data collection process. S.R. contributed to the refinement of the theoretical framework, supervised the validation procedures, and coordinated the analysis of the findings. F.D.M.N. contributed to the data interpretation, manuscript organization, and critical revision of the article to align with international publication standards. All authors reviewed and approved the final version of this manuscript. The percentage contributions for the conceptualization, drafting, and revision of this article are as follows: E.W.S.D.P.: 45%, S.R.: 35%, and F.D.M.N.: 20%.

7. Data Availability Statement

The authors state that the data supporting the findings of this study will be made available by the corresponding author, S.R, upon reasonable request.

REFERENCES

Abulela, M. A. A., et al. (2024). Development and initial validation of a Creative Self-Efficacy (CSE) scale for undergraduates. Frontiers in Education, 9, 1306532. https://doi.org/10.3389/feduc.2023.1306532

Amabile, T. M., & Pratt, M. G. (2016). The dynamic componential model of creativity and innovation in organizations: Making progress, making meaning. Research in Organizational Behavior, 36, 157–183. https://doi.org/10.1016/j.riob.2016.10.001

Beaty, R. E., & Kenett, Y. N. (2023). Associative thinking at the core of creativity. Trends in Cognitive Sciences, 27(7), 671–683. https://doi.org/10.1016/j.tics.2023.04.004

Capron Puozzo, I., & Audrin, C. (2021). Improving self-efficacy and creative self-efficacy to foster creativity and learning in schools. Thinking Skills and Creativity, 42, 100965. https://doi.org/10.1016/j.tsc.2021.100965

Christensen, A. P., Cardillo, E. R., & Chatterjee, A. (2023). Aesthetic cognitivism: A review and an empirical agenda. Psychology of Aesthetics, Creativity, and the Arts, 17(S1), 26–43. https://doi.org/10.1037/aca0000537

Izadi-Avanji, F. S., Zamani, N., Azizi-Fini, I., & Mohammadnejad, E. (2024). Development and psychometric testing of the Aesthetics of Everyday Life Scale in Aging (AELSA). BMC geriatrics, 24(1), 259.

Fekete, A., Pelowski, M., Specker, E., Brieber, D., Rosenberg, R., & Leder, H. (2023). The Vienna Art Picture System (VAPS): A dataset of 999 paintings and subjective ratings for art and aesthetics research. Psychology of Aesthetics, Creativity, and the Arts, 17(5), 660–671. https://doi.org/10.1037/aca0000460

García-Madruga, J. A., Lupiáñez, J., & Navarro, E. (2025). Development and validation of the Scale of Aesthetics and Creativity in Chess (SACC). Frontiers in Psychology, 16, 1456.

Iosifyan, M. (2021). Theory of mind increases aesthetic appreciation in visual arts. Psychology of Aesthetics, Creativity, and the Arts, 15(3), 451–461. https://doi.org/10.1037/aca0000301

Kang, S., & Kim, J. (2017). Evaluating students' performance in responding to art: Development and validation of an Art Criticism Assessment Rubric (ACAR). British Journal of Educational Studies, 65(4), 489–507.

Karwowski, M., Lebuda, I., & Wiśniewska, E. (2024). Development and initial validation of a creative self-efficacy scale for undergraduates. Frontiers in Education, 9, 223.

Li, F., Wang, L., & Sun, J. (2023). Construction and validation of the Postgraduate Research Innovation Ability Scale. Frontiers in Psychology, 14, 445.

Li, F., et al. (2024). Creativity self-assessments in design education: A systematic review. Thinking Skills and Creativity, 52, 101350. https://doi.org/10.1016/j.tsc.2024.101350

Liu, D., Wu, C., Meng, Y., & Dang, J. (2025). How self-belief in creativity and well-being is associated with life satisfaction, meaning in life, and psychological richness. Journal of Intelligence, 13(3), 25. https://doi.org/10.3390/jintelligence13030025

Liu, M., Pei, Y., & Pang, W. (2025). Exploring the Interplay Between Creative Self-Efficacy, Mindset, and Creativity in Response to Negative Feedback. Behavioral Sciences, 15(6), 806.

Mattheiss, J., & Herrmann, H. (2024). Fostering appreciation of beauty in education. Frontiers in Education, 9, 1381101. https://doi.org/10.3389/feduc.2024.1381101

Muth, C., Carbon, C. C., & Leder, H. (2023). Aesthetic Processing Preference Scale (APPS): Development and validation. Psychology of Aesthetics, Creativity, and the Arts, 17(3), 432–446.

Pei, X., & Cao, Y. (2025). Aesthetic appreciation is different: Aesthetic pleasure in architectural aesthetics. Buildings, 15(4), 543. https://doi.org/10.3390/buildings15040543

Pelowski, M., Markey, P. S., Forster, M., Gerger, G., & Leder, H. (2017). Move me, astonish me... delight my eyes and brain: The Vienna Integrated Model of Art Perception (VIMAP) and its application to empirical aesthetics. Physics of Life Reviews, 21, 80–125. https://doi.org/10.1016/j.plrev.2017.02.003

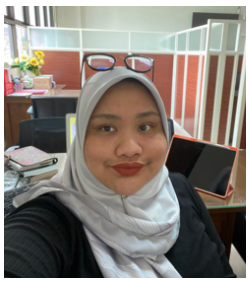
Pizzolante, M., Pelowski, M., Demmer, T. R., Bartolotta, S., Sarcinella, E. D., Gaggioli, A., & Chirico, A. (2024). Aesthetic experiences and their transformative power: A systematic review. Frontiers in Psychology, 15, 1328449. https://doi.org/10.3389/fpsyg.2024.1328449

Primi, C., Benedetti, D., & Donati, M. (2021). Development and validation of an inventory to assess traits of creative potential (TCPQ-12). Frontiers in Psychology, 12, 678901.

Sawada, K., et al. (2024). Appreciation processing evoking feelings of being moved and inspiration: Awe and meaning-making. Journal of Aesthetics & Creativity Research, 2(1). https://doi.org/10.1016/j.jacr.2024.100025

Tao, Y., Chen, S., Li, J., & Zhang, Y. (2024). The influence of trait and state creative self-efficacy on creative behavior. Frontiers in Psychology, 15, 1398257. https://doi.org/10.3389/fpsyg.2024.1398257

Author Biographies

	**E.W Suprihatin Dyah Pratamawati,** is a lecturer and researcher at the Lecturer in the Study Program of Performing Arts Education, Department of Arts and Design, Faculty of Letters, Universitas Negeri Malang. Research expertise in educational development studies. Email: e.w.suprihatin.fs@um.ac.id
	**Dr. Sri Rahayuningsih** is a lecturer and researcher at the department of basic education, faculty of Sekolah Pascasarjana, Universitas Negeri Malang, Jawa Timur, Indonesia. His research interest is Cognitive flexibility Education, Basic mathematics education, and Thinking Processes. Email: srirahayuningsih.pasca@um.ac.id
	**Farah Damia Mohd Nasir** is a researcher at the Faculty of Educational Studies, Universiti Tun Abdul Razak (UNIRAZAK), Malaysia. Her scholarly interests focus on educational research, curriculum innovation, and competency development in 21st-century learning. She has been involved in collaborative research projects related to teacher education, educational assessment, and learning design. Her current research promotes the integration of pedagogical creativity and critical thinking within contemporary education systems. She actively contributes to academic writing, research supervision, and scholarly communication in the field of education. Email: Damia@unirazak.edu.my