

# Designing Authentic Performance Assessment in Primary Music Education: A Framework for Evaluating Musical Creativity and Performance

Joshua Raj Pitchemuthu, Mohd Hassan Abdullah

Sultan Idris Education University, Tanjong Malim, Malaysia.

DOI: <https://doi.org/10.51244/IJRSI.2026.1303000199>

Received: 23 March 2026; Accepted: 28 March 2026; Published: 15 April 2026

## ABSTRACTS

Authentic assessment has increasingly been recognised as an effective approach for evaluating complex learning outcomes in arts education, particularly in music education where learning is demonstrated through performance, creativity, and artistic interpretation. Despite its pedagogical relevance, many primary school teachers encounter difficulties implementing authentic assessment due to the absence of structured frameworks that guide systematic evaluation of musical performance and creative learning. This study aimed to develop and validate an authentic performance assessment framework for primary music education that supports the evaluation of musical creativity and performance-based learning. A Design and Development Research (DDR) approach was employed consisting of three phases: needs analysis, framework development, and expert validation. Survey data were collected from 130 primary school music teachers to examine current assessment practices and challenges. The proposed framework was subsequently validated using the Fuzzy Delphi Method with eight experts in music education and educational assessment. The results indicated strong expert consensus (97%) regarding the relevance and practicality of the framework. The validated model conceptualises authentic music assessment as an integrated cycle involving authentic musical tasks, multidimensional evaluation criteria, structured documentation tools, and reflective feedback mechanisms. The study contributes to music education scholarship by proposing a process-oriented framework that aligns assessment with the experiential and creative nature of musical learning. The framework provides practical guidance for teachers and offers policy implications for strengthening performance-based assessment in primary music education.

**Keywords:** authentic assessment, music education, performance assessment, creativity, primary education

## INTRODUCTION

Assessment in music education presents distinctive pedagogical and methodological challenges due to the inherently multidimensional nature of musical learning. Unlike many academic disciplines in which learning outcomes can be evaluated primarily through written examinations or knowledge-based testing, musical competence is typically demonstrated through performance, interpretation, and creative engagement with musical material. As a result, meaningful evaluation of musical learning requires assessment approaches capable of capturing the integration of cognitive understanding of musical concepts, psychomotor performance skills, affective engagement with artistic expression, and the ability to generate original musical ideas (McPherson & Welch, 2018; Biasutti, 2020).

These characteristics make the assessment of musical learning inherently complex. Conventional assessment practices that emphasise written testing and knowledge recall often fail to capture the breadth and depth of musical competence. Instead, evaluating musical learning requires approaches that reflect the experiential and performative nature of musical activity. Consequently, scholars increasingly advocate the adoption of authentic assessment approaches that allow learners to demonstrate their understanding through tasks that mirror real-world musical practices (Fautley & Daubney, 2019; Ajjawi et al., 2023).

Authentic assessment has emerged as a prominent paradigm in contemporary educational evaluation. Unlike traditional assessment methods, which focus primarily on measuring knowledge acquisition, authentic

assessment emphasises the application of knowledge and skills in meaningful contexts that resemble real disciplinary practices (Gulikers et al., 2018; Wiggins, 2019). Within music education, authentic assessment typically involves performance-based tasks such as singing, instrumental performance, improvisation, composition, and collaborative ensemble participation. These activities provide opportunities for students to demonstrate musical understanding through artistic engagement rather than through abstract theoretical testing.

Recent scholarship suggests that authentic assessment can enhance student engagement, motivation, and deeper learning in arts education contexts. When assessment tasks are closely aligned with meaningful artistic experiences, students are more likely to perceive assessment as an integral component of the learning process rather than as an external judgement imposed by teachers (Panadero et al., 2020; Tai et al., 2022). Authentic assessment therefore supports the development of musical competence and creative thinking by encouraging learners to explore musical ideas through active participation in music-making processes.

Despite these advantages, the implementation of authentic assessment in classroom practice remains challenging. Teachers often encounter difficulties when designing performance-based tasks that align with curriculum objectives while ensuring that assessment decisions remain reliable and transparent. Evaluating musical performance inevitably involves elements of subjective judgement, particularly when explicit performance criteria and structured assessment tools are lacking. Under such circumstances, assessment practices may become inconsistent across classrooms and educational contexts (Fautley & Daubney, 2019).

These challenges are particularly pronounced in primary school settings, where music is frequently taught by generalist teachers who may have limited formal training in music pedagogy or performance assessment. In many education systems, instructional time allocated to music education is relatively limited, and teachers must balance music teaching responsibilities with those of other subjects. Consequently, many teachers report uncertainty regarding how to design and implement systematic assessment strategies capable of evaluating musical creativity and performance in primary classrooms (Colwell & Webster, 2020).

Within the Malaysian education system, the Primary School Standard Curriculum (*Kurikulum Standard Sekolah Rendah*) emphasises holistic student development across cognitive, affective, and psychomotor domains. In music education, this curriculum framework seeks to cultivate musical literacy, aesthetic awareness, creative expression, and performance competence among students. However, empirical research indicates that classroom-based assessment practices in performance-oriented subjects such as music remain uneven across schools, particularly when teachers lack structured guidance for implementing authentic assessment approaches (Mariappan & Osman, 2023).

Although authentic assessment has been widely discussed within music education scholarship, relatively few studies have proposed empirically validated frameworks that support the systematic implementation of authentic assessment in primary education contexts. Much of the existing literature focuses on conceptual discussions of authentic assessment or on individual pedagogical strategies rather than on integrated assessment frameworks that combine authentic tasks, explicit evaluation criteria, structured documentation tools, and reflective feedback mechanisms.

To address this gap, the present study aims to develop and validate an authentic performance assessment framework for primary music education using a Design and Development Research approach. The proposed framework conceptualises music assessment as a cyclical process that integrates authentic musical tasks, multidimensional evaluation criteria, systematic documentation of performance evidence, and reflective feedback processes. By aligning assessment practices with the experiential nature of musical learning, the framework seeks to support teachers in evaluating musical creativity and performance more systematically.

Through this approach, the study contributes to music education scholarship by proposing a theoretically grounded and empirically validated framework that strengthens authentic assessment practices in primary music classrooms.

---

## LITERATURE REVIEW

### **Authentic Assessment in Arts Education**

Authentic assessment has gained increasing attention in contemporary educational research as an approach that evaluates learners' ability to apply knowledge and skills in meaningful and contextually relevant situations. Unlike traditional assessment methods that emphasise the recall of factual knowledge, authentic assessment requires students to demonstrate their understanding through complex tasks that mirror real-world disciplinary practices (Gulikers et al., 2018; Villarroel et al., 2020).

Authentic assessment tasks often involve activities such as collaborative projects, creative productions, problem-solving exercises, and performance-based demonstrations that require learners to apply knowledge in practical contexts. Such tasks are designed not only to measure learning outcomes but also to support learning by encouraging students to engage actively with disciplinary practices (Ajjawi et al., 2023).

The theoretical foundations of authentic assessment are closely associated with constructivist learning theory, which conceptualises learning as an active process in which individuals construct knowledge through interaction with meaningful experiences and social environments (Fosnot, 2018). From a constructivist perspective, assessment should function as an integral component of the learning process by providing opportunities for learners to apply their knowledge, reflect on their performance, and receive feedback that supports further development.

Within arts education, authentic assessment is particularly relevant because artistic learning is fundamentally experiential and performative. Musical learning occurs through active engagement with musical activities such as performing, improvising, composing, and interpreting musical works. These activities require learners to integrate technical skill, expressive interpretation, and creative thinking simultaneously (Burnard, 2021; Biasutti, 2020).

Consequently, scholars argue that meaningful assessment in music education must reflect the performative nature of musical practice. Authentic assessment tasks therefore allow educators to observe how students apply musical knowledge in real artistic contexts while capturing the interpretive and expressive dimensions of musical performance (Fautley & Daubney, 2019).

### **Assessment of Musical Creativity and Performance**

Assessing musical performance presents distinctive challenges because musical competence encompasses multiple interconnected dimensions of learning. Musical performance involves not only technical accuracy but also rhythmic control, tonal quality, expressive interpretation, stylistic understanding, and creative exploration. Evaluating these elements requires assessment approaches capable of capturing both the technical and artistic dimensions of musical activity (McPherson & Welch, 2018).

Historically, assessment practices in music education have often prioritised theoretical knowledge while performance skills are evaluated informally through teacher observation. Although observational assessment can provide valuable insights into students' musical abilities, the absence of explicit evaluation criteria may result in inconsistent assessment practices across classrooms and institutions (Fautley & Daubney, 2019).

To address these challenges, researchers increasingly emphasise the importance of structured evaluation frameworks that incorporate clearly defined performance criteria and systematic scoring procedures. Such frameworks can enhance transparency in assessment by ensuring that evaluation decisions are guided by explicit performance standards rather than by subjective impressions alone (Colwell & Webster, 2020).

One widely recommended approach in performance-based disciplines is the use of analytic rubrics. Analytic rubrics allow educators to evaluate multiple aspects of performance individually while providing detailed feedback that supports student improvement. In music education, rubric-based assessment can facilitate the evaluation of various performance dimensions such as technical accuracy, musical expression, creativity, and communication (Panadero et al., 2020).

At the same time, scholars caution that assessment frameworks must maintain a balance between structured evaluation and creative freedom. Excessive standardisation may restrict artistic exploration by encouraging conformity rather than innovation. Effective music assessment frameworks must therefore provide sufficient structure to support reliable evaluation while preserving opportunities for creative interpretation and expressive experimentation (Burnard, 2021).

### **Teacher Assessment Literacy in Music Education**

The successful implementation of authentic assessment depends heavily on teachers' assessment literacy, which refers to educators' knowledge, skills, and professional judgement related to designing assessment tasks, interpreting assessment evidence, and using assessment information to support student learning (DeLuca et al., 2019).

In the context of music education, assessment literacy involves the ability to evaluate performance-based learning outcomes and creative musical activities. Teachers must possess the pedagogical knowledge required to design authentic musical tasks, establish appropriate evaluation criteria, and provide feedback that supports students' artistic development.

However, research suggests that many teachers encounter difficulties implementing effective music assessment practices due to limited training in both assessment design and music pedagogy. These challenges are particularly evident in primary education contexts, where generalist teachers often teach multiple subjects, including music. Without specialised training in music assessment, teachers may feel uncertain about how to evaluate musical performance, creativity, and expressive interpretation in systematic ways (Colwell & Webster, 2020).

Consequently, there is an increasing need for structured assessment frameworks that can support teachers in implementing authentic assessment practices consistently in classroom settings. Such frameworks can provide practical guidance for designing authentic musical tasks, establishing performance criteria, documenting learning evidence, and delivering meaningful feedback that supports students' artistic growth.

## **CONCEPTUAL FRAMEWORK**

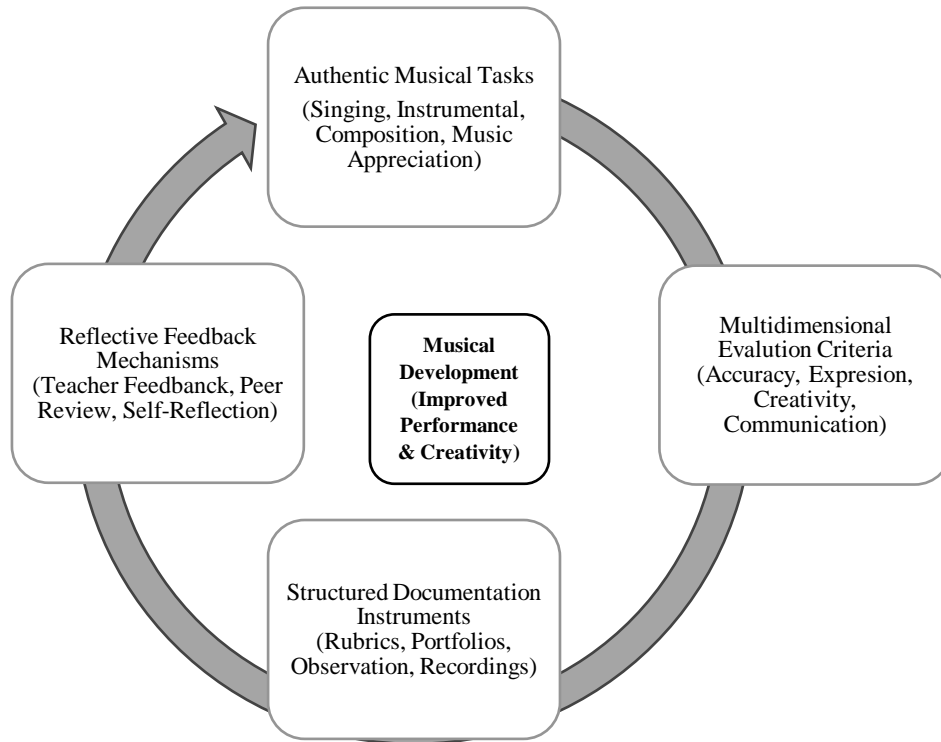
Drawing on the theoretical perspectives discussed in the preceding sections, this study proposes a conceptual framework for authentic performance assessment in primary music education. The framework conceptualises music assessment as a cyclical pedagogical process that integrates performance-based learning activities with systematic evaluation, documentation of learning evidence, and reflective feedback mechanisms. This process-oriented perspective reflects contemporary views of assessment as an integral component of the learning process rather than as a discrete evaluative event (Black & Wiliam, 2018; Tai et al., 2022).

The proposed framework is grounded in two key theoretical traditions: constructivist learning theory and authentic assessment theory. Constructivist perspectives emphasise that meaningful learning occurs when learners actively construct knowledge through engagement with authentic experiences and social interaction (Fosnot, 2018). Within music education, such learning typically occurs through participation in artistic practices such as performing, improvising, composing, and interpreting musical works. Authentic assessment theory similarly emphasises that assessment tasks should reflect real disciplinary practices and provide opportunities for learners to demonstrate knowledge and skills in meaningful contexts (Gulikers et al., 2018; Wiggins, 2019).

Based on these theoretical foundations, the framework conceptualises authentic music assessment as an integrated system composed of four interrelated components:

1. Authentic musical tasks
2. Multidimensional evaluation criteria
3. Structured documentation instruments
4. Reflective feedback mechanisms

These components interact dynamically to support the systematic evaluation of musical learning while maintaining alignment with the experiential and creative nature of music education. The relationships among these four components are illustrated in Figure 1, which conceptualises authentic music assessment as a cyclical process connecting performance, evaluation, documentation, and reflective feedback.



**Figure 1. Authentic Performance Assessment Framework for Primary Music Education**

The framework conceptualises authentic music assessment as a cyclical process in which authentic musical tasks generate performance evidence that is interpreted using multidimensional evaluation criteria, documented through structured assessment instruments, and transformed into reflective feedback that supports continuous musical development.

### **Musical Tasks**

Authentic musical tasks constitute the central component of the framework because they generate observable evidence of students' musical competence. In contrast to traditional assessment approaches that rely primarily on written examinations, authentic tasks require students to demonstrate their understanding through activities that resemble real musical practices. In primary music education, such tasks may include singing, instrumental performance, improvisation, composition, and ensemble participation. These activities allow students to apply musical knowledge and skills within meaningful artistic contexts, thereby providing richer evidence of musical understanding than theoretical tests alone. As noted by McPherson and Welch (2018), musical learning is inherently performative, and therefore assessment strategies must reflect the processes through which musical competence develops in practice. Authentic musical tasks also provide opportunities for teachers to observe how students integrate multiple dimensions of musical learning simultaneously, including technical skill, expressive interpretation, and creative exploration. These activities therefore serve as the primary source of assessment evidence within the proposed framework (Biasutti, 2020; Burnard, 2021).

### **Multidimensional Evaluation Criteria**

While authentic musical tasks generate performance evidence, evaluation criteria provide the interpretive framework through which this evidence is assessed. Because musical performance involves multiple dimensions of competence, effective evaluation requires criteria that capture both technical and artistic aspects of musical expression. Typical evaluation dimensions in music performance assessment include technical accuracy, rhythmic precision, tonal quality, expressive interpretation, creativity, and performance communication. By

defining explicit criteria for these dimensions, educators can evaluate musical performances more systematically and transparently. The use of clearly articulated evaluation criteria is widely recommended in performance-based disciplines because it enhances the reliability and validity of assessment decisions. Explicit criteria help reduce subjectivity in evaluation while also providing learners with clearer expectations regarding performance standards (Fautley & Daubney, 2019; Colwell & Webster, 2020). In addition, structured criteria support the use of analytic rubrics, which enable teachers to evaluate different aspects of performance separately and provide more detailed feedback to students (Panadero et al., 2020).

### **Structured Documentation Instruments**

The third component of the framework involves structured documentation instruments, which enable teachers to systematically record and organise evidence of students' musical learning. Because musical development often occurs gradually through repeated performance experiences, assessment practices must capture evidence of learning over time rather than relying solely on isolated evaluation events. Structured documentation tools may include observation checklists, performance rating forms, student portfolios, reflective journals, and audio or video recordings of musical performances. These tools allow educators to track students' progress across multiple learning activities while providing a comprehensive record of their musical development. Documentation instruments also support formative assessment processes by enabling teachers to analyse patterns in student performance and identify areas requiring further instructional support. In this way, documentation functions not only as a record of learning outcomes but also as a resource for informing subsequent teaching and learning decisions (Black & Wiliam, 2018).

### **Reflective Feedback Mechanisms**

The final component of the framework involves reflective feedback mechanisms, which transform assessment evidence into opportunities for learning improvement. Feedback plays a central role in formative assessment because it enables students to understand the strengths and weaknesses of their performance and identify strategies for improvement (Nicol & Macfarlane-Dick, 2020). Within music education, feedback may take several forms, including teacher feedback, peer feedback, and self-reflection. These processes encourage learners to critically evaluate their own performances and develop greater awareness of musical quality and artistic interpretation. Such reflective practices are closely related to the development of evaluative judgement, which refers to learners' ability to assess the quality of work and make informed decisions about how to enhance it (Tai et al., 2022). By incorporating reflective feedback mechanisms into the assessment process, the proposed framework positions assessment as an iterative learning cycle rather than a terminal judgement of performance. Through repeated cycles of performance, evaluation, documentation, and reflection, students are able to refine their musical skills and deepen their artistic understanding.

### **Integration of Framework Components**

Taken together, these four components form an integrated system for authentic music assessment. Authentic musical tasks generate evidence of performance; evaluation criteria provide interpretive standards for assessing this evidence; documentation instruments capture and organise learning data; and reflective feedback mechanisms enable students and teachers to use assessment information to support further development. By conceptualising assessment as a cyclical interaction between these components, the framework aligns evaluation practices with the experiential and creative nature of musical learning. This integrated approach not only enhances the reliability and transparency of assessment but also supports the broader educational objective of fostering musical creativity, performance competence, and reflective artistic practice in primary education.

## **METHODOLOGY**

### **Research Design**

This study adopted a Design and Development Research (DDR) approach to construct and validate an authentic performance assessment framework for primary music education. DDR is particularly suitable for the development of pedagogical models and educational frameworks because it systematically integrates theoretical

knowledge, empirical investigation, and expert validation in the design of practical educational solutions (Richey & Klein, 2007; McKenney & Reeves, 2019). Unlike purely experimental research designs that focus primarily on hypothesis testing, DDR emphasises the iterative development and refinement of educational innovations grounded in both theory and empirical evidence. The use of DDR in the present study is justified by the need to design an assessment framework that is both conceptually grounded and practically applicable within classroom contexts. Music assessment, particularly in primary education, involves complex performance-based learning outcomes that require structured evaluation tools and systematic implementation procedures. Consequently, a design-oriented research methodology provides an appropriate framework for integrating empirical insights from practitioners with theoretical principles derived from assessment and music education research. The research process was conducted in three sequential phases, each addressing a distinct objective in the development and validation of the proposed framework. Table 1 summarises the structure of the research design.

**Table 1. Design and Development Research Phases**

Phase	Purpose	Method
Phase 1	Identify current challenges and needs in music assessment practices	Survey of primary music teachers
Phase 2	Develop the authentic performance assessment framework	Literature synthesis and curriculum analysis
Phase 3	Validate the framework structure and components	Fuzzy Delphi Method

This multi-phase design enabled the framework to be constructed through a systematic progression from empirical exploration to theoretical development and expert validation, ensuring both conceptual robustness and practical relevance.

### Phase 1: Needs Analysis

#### Participants

The needs analysis phase involved 130 primary school music teachers who were actively engaged in classroom-based music instruction. Participants were drawn from three types of Malaysian national primary schools:

- National Primary Schools (SK)
- National-Type Chinese Primary Schools (SJKC)
- National-Type Tamil Primary Schools (SJKT)

The inclusion of teachers from different school types allowed the study to capture diverse perspectives regarding classroom assessment practices within the Malaysian primary education system. These participants were selected because they are directly responsible for designing and implementing music assessment activities in primary classrooms. Their experiences therefore provide valuable insights into the practical challenges associated with evaluating musical performance and creativity in real educational settings. Sampling teachers across multiple school contexts also enhanced the ecological validity of the findings by ensuring that the needs analysis reflected the varied instructional conditions under which music education is delivered in Malaysian primary schools.

#### Instrument Development

Data for the needs analysis were collected using a structured questionnaire consisting of 55 items designed to measure teachers' perceptions of authentic music assessment practices. The questionnaire employed a five-point Likert scale, ranging from 1 (strongly disagree) to 5 (strongly agree). The instrument was constructed based on key constructs identified in the authentic assessment literature and performance-based evaluation research. Specifically, four conceptual dimensions of authentic music assessment were measured:

1. Authentic performance tasks

2. Evaluation criteria for musical performance
3. Assessment documentation tools
4. Rubric-based performance evaluation

The development of the instrument was informed by existing theoretical frameworks in authentic assessment research, which emphasise the importance of aligning assessment tasks with real disciplinary practices while incorporating explicit evaluation criteria and structured feedback mechanisms (Ajjawi et al., 2023; Gulikers et al., 2018; Wiggins, 2019).

### Reliability Testing

Prior to the main data collection, a pilot study involving 30 primary school music teachers was conducted to assess the reliability of the instrument. Internal consistency reliability was evaluated using Cronbach's alpha coefficient, which measures the extent to which questionnaire items consistently reflect the underlying constructs. The analysis yielded a Cronbach's alpha value of 0.989, indicating excellent internal reliability. According to established methodological guidelines, Cronbach's alpha values above 0.90 reflect very high internal consistency among measurement items (Hair et al., 2019; Taber, 2018). The results therefore suggest that the instrument was sufficiently reliable for use in the main study.

### Exploratory Factor Analysis

To examine the dimensional structure of authentic music assessment practices, Exploratory Factor Analysis (EFA) was conducted using principal component extraction with varimax rotation. EFA is widely recommended in survey-based research to identify latent constructs underlying measurement items and to confirm the conceptual structure of research instruments (Hair et al., 2019). The analysis revealed four distinct factors corresponding to the theoretical constructs of authentic music assessment. Table 2 summarises the factor structure identified through the analysis.

**Table 2. Exploratory Factor Analysis Results**

Factor	Number of Items	Factor Loadings	Variance Explained
Authentic performance tasks	12	.71 – .88	22.6%
Evaluation criteria	14	.69 – .84	19.1%
Assessment documentation	15	.63 – .81	15.7%
Reflective feedback	14	.66 – .83	11.4%

The four factors collectively explained 68.8% of the total variance, indicating a strong conceptual structure underlying authentic assessment practices in music education. Factor loadings across all items exceeded the recommended threshold of 0.60, suggesting that the measurement items demonstrated satisfactory construct validity (Hair et al., 2019).

### Phase 2: Framework Development

The second phase involved the development of the Authentic Performance Assessment Framework (APAF). The framework was constructed through the integration of three complementary sources of evidence:

1. Empirical findings obtained from the needs analysis
2. Theoretical insights derived from authentic assessment literature
3. Curriculum analysis of the Malaysian Primary School Standard Curriculum (KSSR)

This triangulated approach ensured that the framework was simultaneously grounded in empirical teacher perspectives, established theoretical principles, and national curriculum requirements. The resulting framework

conceptualises authentic music assessment as a cyclical process involving four interrelated components. Table 3 summarises the structure of the framework.

**Table 3. Core Components of the Authentic Performance Assessment Framework**

Component	Function
Authentic musical tasks	Generate observable evidence of musical performance
Evaluation criteria	Interpret performance outcomes across multiple musical dimensions
Assessment documentation	Record and monitor students' musical development
Reflective feedback	Support learning improvement through feedback and reflection

This integrated structure reflects contemporary perspectives on assessment as learning, which conceptualise assessment as an ongoing process that supports student development through feedback, reflection, and iterative improvement (Tai et al., 2022).

### Phase 3: Framework Validation

The final phase of the study involved validating the proposed framework using the Fuzzy Delphi Method (FDM). The Delphi technique is widely used in educational research to obtain consensus among experts regarding complex conceptual models or policy recommendations (Okoli & Pawlowski, 2004). The fuzzy extension of the Delphi method enhances the reliability of expert judgement by incorporating fuzzy set theory to address uncertainty and ambiguity in human decision-making (Chang et al., 2011).

### Expert Panel

Eight experts participated in the validation process. The panel was deliberately composed of individuals with diverse expertise relevant to music education, curriculum development, and educational assessment. Table 4 presents the composition of the expert panel.

**Table 4. Expert Panel Composition**

Expert	Area of Expertise
Music education professor	Curriculum design
Educational assessment specialist	Measurement theory
Music pedagogy researcher	Creativity education
Teacher training expert	Music pedagogy
Curriculum officer	National curriculum
Senior music educator	Classroom practice
Educational psychologist	Learning theory
Assessment researcher	Evaluation design

The inclusion of experts from multiple disciplinary perspectives ensured that the framework was evaluated from both theoretical and practical viewpoints.

### Delphi Instrument

Experts evaluated 32 elements of the proposed framework using a five-point Likert scale ranging from:

1 = Strongly disagree

5 = Strongly agree

The evaluation criteria focused on three key dimensions:

- relevance of framework components
- clarity of conceptual structure
- practicality for classroom implementation

### Fuzzy Delphi Analysis

The validation analysis involved three key indicators commonly used in Fuzzy Delphi studies:

1. Threshold value (d) – measures the degree of agreement between expert opinions
2. Expert consensus percentage – indicates the proportion of experts supporting each framework element
3. Defuzzification score – converts fuzzy values into a single crisp value representing the final consensus score

**Table 5. Fuzzy Delphi Validation Results**

Indicator	Result	Acceptance Criterion
Threshold value (d)	0.015	$\leq 0.20$
Expert consensus	97%	$\geq 75\%$
Defuzzification score	0.78	$\geq 0.50$

The threshold value obtained in this study ( $d = 0.015$ ) was substantially lower than the maximum acceptable value of 0.20, indicating a high level of agreement among experts. Similarly, the expert consensus level of 97% exceeded the minimum threshold of 75%, demonstrating strong agreement regarding the framework components. The defuzzification score of 0.78 further confirmed that the proposed framework elements were considered appropriate and relevant for implementation in primary music education contexts. Taken together, these results demonstrate strong expert consensus regarding the conceptual validity and practical applicability of the Authentic Performance Assessment Framework.

## FINDINGS

### Teachers’ Perceptions of Authentic Music Assessment

The results of the needs analysis survey indicate a strong level of teacher support for the adoption of authentic assessment practices in primary music education. Descriptive statistical analysis revealed that the mean scores for the four key constructs of authentic music assessment ranged from 4.13 to 4.29 on a five-point Likert scale, reflecting a consistently high level of agreement among participating teachers regarding the importance of structured assessment approaches that align with the experiential nature of music learning.

**Table 6. Teachers’ Perceptions of Authentic Music Assessment**

Assessment Component	Mean	Interpretation
Assessment strategies	4.29	Very high
Evaluation criteria	4.21	Very high
Assessment instruments	4.18	High
Rubric-based assessment	4.13	High

Among the four constructs, assessment strategies recorded the highest mean score ( $M = 4.29$ ), indicating that teachers place strong emphasis on the use of authentic performance-based tasks as a central component of music assessment. This finding reflects the experiential and performative nature of musical learning, where students demonstrate understanding primarily through practical engagement rather than through theoretical examinations. The high level of agreement observed across all constructs suggests that teachers recognise the importance of systematic assessment frameworks that integrate authentic learning activities with clearly defined evaluation criteria and structured documentation tools. These results align with previous research emphasising that effective music assessment requires the alignment of performance tasks, assessment criteria, and evaluation instruments in order to capture the multidimensional nature of musical competence (Fautley & Daubney, 2019; Colwell & Webster, 2020). Furthermore, the findings suggest that teachers perceive analytic rubrics and structured assessment tools as essential for improving transparency and consistency in performance evaluation. Rubric-based assessment has been widely recommended in arts education because it allows educators to evaluate multiple dimensions of performance while providing explicit performance expectations for learners (Panadero et al., 2020). Consequently, the strong support for rubric-based evaluation observed in this study highlights teachers' recognition of the need for systematic assessment approaches capable of capturing the technical, expressive, and creative aspects of musical learning.

### **Performance-Based Assessment Practices in Music Education**

The survey findings further indicate that teachers strongly support the use of performance-oriented assessment strategies in music education. Respondents reported that musical learning outcomes are most effectively evaluated through authentic musical activities that require students to demonstrate their knowledge and skills through active participation in music-making processes.

The most frequently identified performance-based assessment activities include:

- vocal performance (singing)
- instrumental performance
- ensemble participation
- composition or creative music projects

These activities correspond closely with the core processes through which musical competence is developed and demonstrated in educational settings. Performance-based tasks enable teachers to observe not only students' technical proficiency but also their expressive interpretation, creative engagement, and collaborative musical interaction. The emphasis on performance-oriented assessment identified in this study is consistent with contemporary perspectives in music education research, which emphasise that musical competence is best evaluated through authentic artistic practices rather than through written theoretical tests (Biasutti, 2020; McPherson & Welch, 2018). Performance activities provide rich sources of assessment evidence because they allow educators to observe how students apply musical knowledge in real musical contexts. Moreover, authentic musical activities such as ensemble participation and creative composition tasks can reveal dimensions of learning that may not be captured through conventional assessment methods. These activities require students to integrate technical skills, interpretive judgement, and creative thinking simultaneously, thereby providing a more holistic representation of musical understanding (Burnard, 2021; Hickey, 2022). The findings therefore reinforce the argument that music assessment frameworks should prioritise performance-based learning activities while incorporating structured evaluation mechanisms to ensure reliable interpretation of performance outcomes.

### **Expert Validation of the Authentic Performance Assessment Framework**

Following the development of the proposed Authentic Performance Assessment Framework (APAF), the model was evaluated through the Fuzzy Delphi Method (FDM) to determine the degree of expert consensus regarding its conceptual relevance and practical applicability. Three key indicators were used to evaluate the level of agreement among the expert panel: threshold value ( $d$ ), percentage of expert consensus, and defuzzification score. Table 7 summarises the results of the validation analysis.

**Table 7. Fuzzy Delphi Validation Results**

Indicator	Result	Acceptance Criterion
Threshold value (d)	0.015	$\leq 0.20$
Expert consensus	97%	$\geq 75\%$
Defuzzification score	0.78	$\geq 0.50$

The threshold value obtained in this study ( $d = 0.015$ ) is substantially lower than the maximum acceptable value of 0.20, indicating a very high degree of agreement among experts regarding the proposed framework elements. According to Chang et al. (2011), lower threshold values reflect stronger consensus in Delphi-based evaluations because they indicate minimal divergence in expert judgements. Similarly, the expert consensus level of 97% far exceeds the minimum acceptance threshold of 75%, demonstrating broad agreement among experts regarding the relevance and conceptual clarity of the framework components. The defuzzification score of 0.78, which is well above the minimum requirement of 0.50, further confirms that the framework elements were considered appropriate for implementation within primary music education contexts. Taken together, these results provide strong evidence supporting the conceptual validity of the Authentic Performance Assessment Framework. The high level of expert consensus suggests that the integration of authentic musical tasks, multidimensional evaluation criteria, structured documentation tools, and reflective feedback processes represents a coherent and practically applicable model for evaluating musical creativity and performance in primary classrooms. These findings are consistent with previous educational design research demonstrating that Delphi-based expert validation can effectively support the development of pedagogical frameworks by ensuring that proposed models reflect both theoretical soundness and professional expertise (Hsu & Sandford, 2007; Okoli & Pawlowski, 2004).

### **Pilot Implementation of the Framework**

To enhance the practical applicability of the proposed framework, a small-scale pilot implementation was conducted in two primary school classrooms over a four-week period. The framework was applied in teaching units involving singing and group instrumental performance. Teachers utilised the framework by designing authentic musical tasks, applying analytic rubrics based on multidimensional evaluation criteria, documenting student performance using observation checklists and video recordings, and providing structured reflective feedback. Findings from the pilot implementation indicated that:

1. teachers were able to apply the framework with moderate ease after initial familiarisation,
2. students demonstrated increased engagement and participation during performance-based activities, and
3. rubric-based evaluation improved transparency in assessment.

However, several challenges were identified, including time constraints during lesson implementation, increased teacher workload in documentation, and the need for training in rubric design and feedback strategies. These findings suggest that while the framework is practically applicable, its effectiveness depends on adequate teacher preparation and contextual adaptation.

## **DISCUSSIONS**

The findings of this study offer several important insights into the implementation of authentic assessment practices in primary music education. By integrating empirical evidence from teachers with expert validation through the Fuzzy Delphi Method, the study contributes to ongoing discussions concerning how assessment frameworks can more effectively capture the experiential and creative nature of musical learning. The discussion that follows interprets the findings in relation to contemporary theoretical perspectives on authentic assessment, performance-based learning, and assessment literacy in arts education.

First, the results indicate that teachers strongly recognise the importance of performance-based assessment strategies as a central component of music education evaluation. This finding reinforces constructivist learning theory, which emphasises that meaningful learning occurs when learners actively construct knowledge through engagement with authentic experiences (Fosnot, 2018; McKenney & Reeves, 2019). In music education,

learning is inherently experiential because musical understanding emerges through participation in activities such as singing, instrumental performance, improvisation, and composition. Consequently, assessment practices that focus on practical musical engagement are better positioned to capture the multidimensional nature of musical competence. The emphasis on performance-based assessment observed in this study aligns with a growing body of research indicating that musical knowledge cannot be adequately assessed through written examinations alone. Instead, musical competence is most effectively demonstrated through artistic practice, where technical skill, expressive interpretation, and creative thinking interact simultaneously (Biasutti, 2020; McPherson & Welch, 2018). Performance-based tasks therefore provide authentic contexts in which students can demonstrate their musical understanding in ways that reflect real musical practices. These findings support arguments in music education scholarship that authentic musical activities represent a more valid form of assessment evidence than traditional academic testing methods (Fautley & Daubney, 2019).

Second, the findings highlight the importance of structured evaluation criteria in ensuring the reliability and transparency of performance assessment. While authentic tasks provide opportunities for students to demonstrate musical competence, the evaluation of artistic performance inevitably involves elements of subjective judgement. Without clearly defined performance criteria, assessment decisions may become inconsistent and difficult to interpret across different classrooms or teachers. Previous research has emphasised that authentic assessment requires the alignment of performance tasks with explicit evaluation standards in order to produce meaningful and reliable assessment outcomes (Ajjawi et al., 2023; Wiggins, 2019). The strong teacher support for structured evaluation criteria identified in this study therefore reflects a growing recognition that performance-based learning must be accompanied by systematic evaluation frameworks. Explicit criteria enable teachers to assess complex musical performances more consistently while also providing students with clearer expectations regarding performance quality. Moreover, clearly articulated criteria facilitate the provision of constructive feedback that supports students' ongoing artistic development.

Third, the findings demonstrate that teachers place considerable value on the use of structured assessment documentation tools, including observation checklists, performance recordings, portfolios, and analytic rubrics. These tools allow teachers to capture evidence of musical learning over extended periods of time, thereby providing a more comprehensive representation of students' musical development. Unlike traditional assessment models that rely on single evaluation events, performance-based disciplines such as music require assessment practices capable of documenting gradual improvement in technical skill, expressive interpretation, and creative thinking. The use of documentation tools also supports formative assessment processes, which emphasise feedback and reflection as central mechanisms for improving learning outcomes. Formative assessment has been widely recognised as a powerful approach for enhancing student learning because it enables teachers to identify learning gaps and provide targeted feedback during the learning process rather than solely at the end of instruction (Black & Wiliam, 2018; Nicol & Macfarlane-Dick, 2020). Within music education, formative assessment practices can encourage students to reflect on their own performance, evaluate their artistic decisions, and refine their musical interpretations through iterative practice.

Another important finding of this study concerns the conceptual validity of the proposed Authentic Performance Assessment Framework, which was confirmed through expert evaluation using the Fuzzy Delphi Method. The high level of expert consensus obtained in the validation process indicates strong agreement regarding the relevance, clarity, and practicality of the framework components. This validation process strengthens the credibility of the framework by demonstrating that it reflects both theoretical insights and professional expertise within the field of music education.

The framework developed in this study integrates four essential components: authentic musical tasks, multidimensional evaluation criteria, structured assessment documentation, and reflective feedback mechanisms. Together, these elements form a coherent assessment system that aligns with the experiential nature of musical learning. Rather than conceptualising assessment as a single evaluative event, the framework positions assessment as an ongoing pedagogical process that supports continuous musical development.

This conceptualisation extends existing literature on authentic assessment in arts education. While previous studies have highlighted the importance of authentic learning tasks in creative disciplines (Gulikers et al., 2018; Wiggins, 2019), the findings of the present study suggest that effective music assessment requires the integration

of several interdependent elements. Authentic performance tasks alone are insufficient without structured evaluation frameworks capable of guiding the interpretation of complex artistic outcomes. Without such frameworks, assessment practices may remain overly dependent on individual teacher judgement and may therefore lack consistency and transparency.

The proposed framework therefore contributes to contemporary discussions on assessment for learning by conceptualising music assessment as a cyclical process that integrates performance, evaluation, documentation, and reflection. This perspective aligns with recent scholarship emphasising the importance of evaluative judgement, which refers to learners' ability to assess the quality of their own work and make informed decisions about how to improve it (Tai et al., 2022). By incorporating reflective feedback processes into the assessment framework, the model encourages students to engage actively in evaluating their own musical development.

Furthermore, the framework addresses an important gap in primary music education research. Although authentic assessment has been widely advocated in arts education literature, relatively few studies have proposed empirically validated frameworks that guide systematic implementation of authentic assessment practices in classroom contexts. Much of the existing research focuses on pedagogical strategies or theoretical discussions rather than on structured assessment models that can be readily applied by teachers.

By integrating empirical insights from classroom practitioners with expert validation, the present study offers a framework that is both theoretically grounded and practically applicable. The Authentic Performance Assessment Framework therefore provides a structured approach for evaluating musical creativity and performance in primary education while maintaining the flexibility necessary to support artistic exploration and creative expression. Despite its strengths, the implementation of the framework may present several practical challenges. First, time constraints within primary school timetables may limit opportunities for conducting detailed performance assessments and reflective feedback sessions. Second, the use of structured documentation tools such as rubrics, portfolios, and recordings may increase teachers' workload, particularly for generalist teachers with limited music training. Third, variations in school resources, including access to musical instruments and recording technologies, may affect the consistency of implementation across different educational contexts. Addressing these challenges requires institutional support, including reduced assessment burden, access to teaching resources, and targeted professional development programmes.

## **Contributions And Implications**

### **Theoretical Contribution**

This study contributes to the growing body of scholarship on authentic assessment in arts education by proposing a process-oriented conceptualisation of music assessment that positions evaluation as an integrated cycle of performance, interpretation, documentation, and reflective learning. While existing literature has widely acknowledged the importance of authentic tasks in performance-based disciplines, much of the prior research has focused primarily on task design rather than on the broader assessment processes through which learning evidence is interpreted and utilised. The Authentic Performance Assessment Framework developed in this study extends existing conceptualisations of authentic assessment by emphasising the dynamic interaction between assessment components rather than treating them as isolated elements. In particular, the framework conceptualises music assessment as a continuous pedagogical cycle in which authentic musical tasks generate performance evidence that is interpreted through multidimensional evaluation criteria, documented through structured assessment tools, and subsequently transformed into reflective feedback that supports further learning. This cyclical conceptualisation aligns with contemporary theoretical perspectives on assessment for learning, which emphasise that assessment should function as an ongoing learning process rather than merely as a mechanism for measuring achievement (Black & Wiliam, 2018; Wiliam, 2018). In addition, the framework resonates with recent scholarship on evaluative judgement, which highlights the importance of enabling learners to develop the capacity to assess the quality of their own work and make informed decisions about how to improve it (Tai et al., 2022). By conceptualising music assessment as a process that integrates performance, evaluation, documentation, and reflection, the present study contributes a theoretical model that more accurately reflects the experiential and creative nature of musical learning. In doing so, the study advances theoretical discussions in music education by demonstrating that authentic assessment in performance-based disciplines

must be understood not only as task authenticity but also as systemic alignment between assessment practices and artistic learning processes.

### **Methodological Contribution**

The study also contributes methodologically to research on educational assessment by demonstrating how Design and Development Research (DDR) can be effectively integrated with Fuzzy Delphi validation to develop and refine pedagogical frameworks within arts education contexts. DDR has increasingly been recognised as a valuable methodological approach for designing educational innovations because it combines theoretical analysis with empirical investigation and iterative validation (McKenney & Reeves, 2019; Richey & Klein, 2007). However, its application in music education research remains relatively limited. By employing a multi-phase DDR design, the present study illustrates how empirical insights from classroom practitioners can be systematically integrated with theoretical knowledge and expert consensus in order to develop assessment frameworks that are both conceptually robust and practically applicable. The inclusion of the Fuzzy Delphi Method further strengthens the methodological rigour of the study by providing a structured mechanism for synthesising expert judgements and resolving ambiguity in the evaluation of complex conceptual models (Chang et al., 2011). This methodological integration demonstrates the potential of combining design-based research approaches with consensus-building techniques in the development of educational assessment frameworks. Such methodological strategies are particularly valuable in performance-based disciplines such as music education, where assessment models must accommodate the inherent complexity and subjectivity of artistic performance.

### **Practical Contribution**

From a practical perspective, the Authentic Performance Assessment Framework provides music educators with a structured approach for evaluating musical creativity and performance in primary classrooms. One of the persistent challenges in music education assessment lies in balancing the need for reliable evaluation with the need to preserve opportunities for artistic exploration and creative expression. Without clear assessment structures, teachers may rely heavily on informal observation, which can lead to inconsistencies in evaluation practices across different educational contexts (Fautley & Daubney, 2019). The framework developed in this study addresses this challenge by integrating authentic performance tasks with clearly articulated evaluation criteria, systematic documentation tools, and reflective feedback processes. By structuring assessment practices around these interconnected components, the framework enables teachers to evaluate musical learning more transparently while still allowing flexibility for creative interpretation. Moreover, the framework supports the implementation of competency-based curricula, which increasingly emphasise the development of higher-order skills such as creativity, collaboration, and critical thinking. In music education, such competencies are often demonstrated through artistic practice rather than through traditional academic testing. Consequently, structured performance-based assessment frameworks can assist teachers in aligning classroom evaluation practices with broader educational goals related to creative and holistic learning (Burnard, 2021; Colwell & Webster, 2020).

### **Implications**

#### **Implications for Music Education Practice**

The findings of this study have several implications for classroom practice in music education. First, the results highlight the importance of aligning assessment practices with the experiential nature of musical learning. Because musical understanding is often expressed through performance and creative engagement, assessment strategies should prioritise authentic musical activities such as singing, improvisation, instrumental performance, and collaborative ensemble participation. The proposed framework supports this alignment by providing teachers with structured guidance for designing and evaluating performance-based learning tasks. By integrating authentic musical activities with clear evaluation criteria and systematic documentation procedures, the framework can enhance both the transparency and consistency of assessment practices in music classrooms. Furthermore, the inclusion of reflective feedback mechanisms within the framework encourages students to engage in metacognitive reflection on their own musical performance, thereby supporting the development of evaluative judgement and self-regulated learning. Research has shown that such reflective assessment practices

can significantly enhance student engagement and learning outcomes in creative disciplines (Panadero et al., 2020; Tai et al., 2022).

### Policy Implications

Beyond classroom practice, the findings of this study also have implications for curriculum development and educational policy. In many educational systems, assessment frameworks in arts education remain heavily influenced by traditional academic evaluation models that prioritise theoretical knowledge over practical competence. However, the findings of this study reinforce the argument that performance-based disciplines require assessment systems capable of capturing the experiential and creative dimensions of learning.

Educational authorities may therefore consider incorporating authentic assessment frameworks into official curriculum guidelines in order to support more consistent evaluation of performance-based learning outcomes in music education. Such frameworks can provide teachers with clearer guidance regarding how musical performance, creativity, and expressive interpretation should be assessed within classroom contexts. In addition, teacher professional development programmes should place greater emphasis on assessment literacy in arts education. Teachers require specialised knowledge and skills in order to design authentic performance tasks, interpret assessment evidence, and provide constructive feedback that supports artistic development. Strengthening teachers' assessment literacy can therefore play a critical role in improving the quality and consistency of music assessment practices (DeLuca et al., 2019).

### Teacher Training Implication

The findings highlight the importance of strengthening teachers' assessment literacy in music education. Professional development programmes should focus on training teachers to design authentic performance tasks, develop analytic rubrics, and provide effective formative feedback. Such training is essential to ensure the successful implementation of the proposed framework in real classroom contexts.

### Future Research Directions

While the present study provides a validated framework for authentic music assessment, several avenues

remain for future research. First, further studies should examine the implementation of the framework in real classroom environments through **longitudinal intervention studies**. Such research could investigate how the framework influences students' musical development, creative thinking, and performance competence over extended periods of instruction. Second, future research may explore how **digital technologies** can enhance authentic music assessment practices.

Emerging digital tools such as electronic portfolios, online performance recording platforms, and artificial intelligence–assisted performance analysis systems offer new possibilities for documenting and evaluating musical learning. These technologies can enable teachers to capture richer evidence of student performance while also facilitating more detailed feedback and reflection processes. Finally, comparative studies across different educational contexts may provide further insights into how authentic assessment frameworks can be adapted to diverse music education systems. Such research would contribute to the development of globally relevant models of performance-based assessment that support meaningful evaluation of musical creativity and artistic expression.

### Limitations of the Study

This study is limited by its reliance on teacher perceptions within the Malaysian primary education context. Although participants were drawn from multiple school types (National Schools, National-Type Chinese Primary Schools and National-Type Tamil Primary School), the findings may not fully represent diverse international educational settings. Future research should validate the framework across different cultural and institutional contexts to enhance its generalisability. Additionally, the framework validation relied primarily on expert consensus rather than large-scale classroom implementation, which may limit conclusions regarding its long-term effectiveness.

---

## CONCLUSION

This study set out to develop and validate an authentic performance assessment framework designed to support the systematic evaluation of musical creativity and performance in primary music education. Drawing on a Design and Development Research (DDR) approach, the study integrated empirical insights from practicing music teachers with expert consensus obtained through the Fuzzy Delphi Method in order to construct a conceptually grounded and practically applicable assessment model.

The findings demonstrate that teachers strongly recognise the need for assessment approaches that reflect the experiential and performance-based nature of musical learning. The results of the needs analysis highlight the importance of authentic musical activities, clearly articulated evaluation criteria, structured documentation tools, and rubric-based evaluation procedures in supporting meaningful assessment of musical competence. These findings reinforce the view that musical learning cannot be adequately evaluated through traditional academic testing alone but instead requires assessment strategies that capture the multidimensional processes through which musical knowledge, technical skill, and creative expression develop.

Building on these insights, the study proposes the Authentic Performance Assessment Framework, which conceptualises music assessment as a cyclical pedagogical process involving authentic musical tasks, multidimensional performance evaluation, systematic documentation of learning evidence, and reflective feedback mechanisms. By integrating these elements within a coherent structure, the framework positions assessment not merely as an evaluative event but as an integral component of the learning process that supports ongoing musical development. In doing so, the framework aligns with contemporary perspectives on assessment for learning, which emphasise feedback, reflection, and iterative improvement as key mechanisms for enhancing student learning outcomes (Black & Wiliam, 2018; Tai et al., 2022).

From a broader perspective, the study contributes to music education scholarship by advancing a process-oriented conceptualisation of authentic assessment that reflects the artistic and experiential characteristics of musical learning. While previous research has emphasised the value of authentic tasks in arts education, the present study demonstrates that effective music assessment requires the systematic integration of task design, evaluation criteria, documentation tools, and reflective feedback processes. Such an integrated approach can enhance the transparency, consistency, and educational value of assessment practices in performance-based disciplines.

In practical terms, the framework offers music educators a structured model for implementing authentic assessment within primary classroom contexts. By providing teachers with clear guidelines for designing performance tasks, evaluating musical outcomes, and documenting student progress, the framework has the potential to support more consistent and meaningful assessment practices that align with competency-based educational goals.

Nevertheless, several avenues remain for further investigation. Future research should examine the implementation of the framework in real classroom environments through longitudinal or intervention-based studies in order to evaluate its impact on students' musical development and creative learning processes. In addition, emerging digital technologies—including electronic portfolios, performance recording platforms, and artificial intelligence–assisted analysis tools—offer promising opportunities for enhancing the documentation and interpretation of musical performance. Exploring how such technologies can be integrated into authentic music assessment frameworks represents an important direction for future research.

In conclusion, this study provides both a conceptual and practical contribution to the field of music education by proposing a validated framework for authentic performance assessment in primary music classrooms. By aligning assessment practices with the experiential nature of musical learning, the framework offers a pathway toward more meaningful evaluation of musical creativity, performance competence, and artistic expression in contemporary music education.

## ACKNOWLEDGEMENTS

The authors express their sincere appreciation to the Faculty of Music and Performing Arts, Sultan Idris Education University (UPSI), for their institutional support and academic guidance. The authors also extend their gratitude to the experts and primary school music teachers who contributed to this study through their valuable insights and participation.

## REFERENCES

1. Ajjawi, R., Tai, J., Dawson, P., Boud, D., & Bearman, M. (2023). From authentic assessment to authenticity in assessment: Broadening perspectives. *Assessment & Evaluation in Higher Education*, 49(4), 499–510. <https://doi.org/10.1080/02602938.2023.2186751>
2. Biasutti, M. (2020). Assessing musical creativity: A systematic review. *Music Education Research*, 22(2), 154–172. <https://doi.org/10.1080/14613808.2019.1703921>
3. Black, P., & Wiliam, D. (2018). Classroom assessment and pedagogy. *Assessment in Education: Principles, Policy & Practice*, 25(6), 551–575. <https://doi.org/10.1080/0969594X.2018.1441807>
4. Burnard, P. (2021). Creativity and music education: Research, practice and policy. In C. Randles & P. Burnard (Eds.), *The Routledge companion to creativities in music education* (pp. 13–28). Routledge.
5. Chang, P., Hsu, C., & Chang, P. (2011). Fuzzy Delphi method for evaluating educational frameworks. *Expert Systems with Applications*, 38(12), 15091–15099. <https://doi.org/10.1016/j.eswa.2011.05.045>
6. Colwell, R., & Webster, P. (2020). *MENC handbook of research on music learning* (Vols. 1–2). Oxford University Press.
7. DeLuca, C., Coombs, A., & LaPointe-McEwan, D. (2019). Assessment literacy in teacher education: A national survey of teacher educators' practices and perspectives. *Teaching and Teacher Education*, 83, 1–11. <https://doi.org/10.1016/j.tate.2019.03.010>
8. Fautley, M., & Daubney, A. (2019). *Assessment in music education*. Oxford University Press.
9. Fosnot, C. T. (2018). *Constructivism: Theory, perspectives, and practice* (2nd ed.). Teachers College Press.
10. Gulikers, J., Bastiaens, T., & Kirschner, P. (2018). Authentic assessment revisited: A critical review of theoretical and practical issues. *Educational Technology Research and Development*, 66(5), 1–21. <https://doi.org/10.1007/s11423-018-9608-3>
11. Hsu, C. C., & Sandford, B. A. (2007). The Delphi technique: Making sense of consensus. *Practical Assessment, Research & Evaluation*, 12(10), 1–8. <https://doi.org/10.7275/pdz9-th90>
12. Mariappan, V., & Osman, K. (2023). Challenges in implementing classroom-based assessment in Malaysia. *International Journal of Instruction*, 16(2), 101–118. <https://doi.org/10.29333/iji.2023.1626a>
13. McKenney, S., & Reeves, T. C. (2019). *Conducting educational design research* (2nd ed.). Routledge.
14. McPherson, G. E., & Welch, G. F. (2018). *Music learning and teaching*. Oxford University Press.
15. Nicol, D. J., & Macfarlane-Dick, D. (2020). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, 45(2), 1–14. <https://doi.org/10.1080/03075079.2018.1541171>
16. Okoli, C., & Pawlowski, S. D. (2004). The Delphi method as a research tool: An example, design considerations and applications. *Information & Management*, 42(1), 15–29. <https://doi.org/10.1016/j.im.2003.11.002>
17. Panadero, E., Andrade, H., & Brookhart, S. (2020). Fostering self-regulated learning through rubrics. *Educational Research Review*, 29, 100301. <https://doi.org/10.1016/j.edurev.2019.100301>
18. Richey, R. C., & Klein, J. D. (2007). *Design and development research: Methods, strategies, and issues*. Routledge.
19. Tai, J., Ajjawi, R., Boud, D., Dawson, P., & Panadero, E. (2022). Developing evaluative judgement: Enabling students to make decisions about the quality of work. *Higher Education Research & Development*, 41(4), 1–15. <https://doi.org/10.1080/07294360.2021.1889019>

20. Taber, K. S. (2018). The use of Cronbach’s alpha when developing and reporting research instruments in science education. *Research in Science Education*, 48, 1273–1296. <https://doi.org/10.1007/s11165-016-9602-2>
21. Villarroel, V., Bloxham, S., Bruna, D., Bruna, C., & Herrera-Seda, C. (2020). Authentic assessment: Creating a blueprint for course design. *Assessment & Evaluation in Higher Education*, 45(5), 1–15. <https://doi.org/10.1080/02602938.2019.1663414>
22. Wiggins, G. (2019). *Educative assessment: Designing assessments to inform and improve student performance*. Jossey-Bass.
23. Wiliam, D. (2018). *Embedded formative assessment (2nd ed.)*. Solution Tree Press.

## Appendix A

**Table 1. Sample Analytic Rubric for Primary Music Performance Assessment**

Criteria	Excellent (4)	Good (3)	Basic (2)	Limited (1)
Technical Accuracy	Accurate pitch and rhythm throughout	Minor errors	Frequent errors	Inaccurate performance
Expression	Highly expressive and musical	Some expressive elements	Limited expression	No expressive quality
Creativity	Original and creative interpretation	Some creative elements	Minimal creativity	No creativity
Communication	Strong audience engagement	Moderate engagement	Limited engagement	No engagement

This rubric illustrates how multidimensional evaluation criteria can be operationalised in classroom assessment practice.

## Appendix B

### Observation Checklist for Authentic Music Performance Assessment

This appendix presents an example of an observation checklist used to document students’ musical performance during authentic assessment activities in primary music education. The instrument is designed to support systematic observation and facilitate consistent evaluation of key performance dimensions.

#### Instructions:

Teachers are required to observe students during performance activities and indicate the appropriate response for each criterion. The remarks column may be used to provide qualitative feedback.

No.	Assessment Dimension	Indicator	Yes (✓)	No (X)	Remarks
1	Rhythm Accuracy	Demonstrates correct rhythmic patterns and maintains steady tempo			
2	Pitch Accuracy	Maintains accurate pitch throughout the performance			
3	Expressive Dynamics	Applies appropriate dynamics (e.g., phrasing, articulation, volume)			
4	Performance Engagement	Participates actively and confidently in individual or group performance			
5	Musical Creativity	Demonstrates originality or creative interpretation			
6	Reflective Ability	Engages in self-reflection or peer feedback after performance			

### **Purpose of the Instrument:**

This checklist serves as a structured documentation tool to capture observable aspects of musical performance in real classroom contexts. It supports formative assessment by enabling teachers to record student progress, provide targeted feedback, and ensure greater consistency and transparency in evaluation practices.

The instrument aligns with the Authentic Performance Assessment Framework by integrating multidimensional evaluation criteria with systematic documentation processes, thereby facilitating meaningful assessment of musical creativity and performance.