

Mapping formative assessment in higher music education: A scoping review of its implementation and impact on musical intelligence

Research Studies in Music Education

1–22

© The Author(s) 2026



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/1321103X251400571

journals.sagepub.com/home/rsm**Le-Xuan Zhang** 

The Education University of Hong Kong, Hong Kong

Zi Yan 

The Education University of Hong Kong, Hong Kong

Xiang Wang

Hong Kong Baptist University, Hong Kong

Abstract

Formative assessment plays a crucial role in education by enabling teachers and students to gather detailed information about students' learning progress and provide feedback to enhance learning. Despite growing interest in formative assessment across disciplines, its application and impact in higher music education remain underexplored. This scoping review addresses this gap by systematically mapping the landscape of formative assessment practices in university-level music education. Guided by a scoping review framework, 22 empirical studies published between 1976 and 2023 were analyzed. Drawing on Sternberg's theory of musical intelligence, the review identifies three primary formative assessment strategies—teacher-directed, peer assessment, and self-assessment—which respectively contribute to the development of students' creative, practical, analytical, and wisdom-based skills. These skills, when integrated, support the comprehensive growth of musical intelligence. Analytical and practical skills were most frequently enhanced through formative assessment. The review also discusses several contextual and institutional challenges, including time constraints, students' readiness for peer and self-evaluation, and teachers' preparedness to implement formative assessment. This study contributes to the field of both educational assessment and music education and offers critical reflections and directions for using formative assessment more effectively to support students' musical learning experiences.

Keywords

formative assessment, music education, musical intelligence, scoping review, student-centered education

Corresponding author:

Le-Xuan Zhang, The Education University of Hong Kong, 10 Lo Ping Road, Tai Po, New Territory, Hong Kong.

Email: zlexuan@eduhk.hk

Introduction

Educational assessment reform has long emphasized the shift from teacher-centered teaching to student-centered learning, with assessment intended to promote learning rather than merely evaluate outcomes (Andrade & Heritage, 2017). Formative assessment addresses this need by providing tools to track and support students' learning progress, offering significant benefits beyond summative evaluations (P. P. Chen & Bonner, 2019). In higher music education, formative assessment is increasingly recognized for its role in enhancing musical performance (Boucher et al., 2021; Liu et al., 2021). However, despite the expansion of formative assessment activities in music education, no review study has yet provided a holistic understanding of the implementation of formative assessment and its impact on music learning. As music learning extends beyond performance, encompassing a range of cognitive and creative skills such as composing, critiquing, and analytical reasoning (Conway, 2020; Kokotsaki & Hallam, 2007), it is important to understand how formative assessment is applied to support the development of these diverse skills. In addition, university-level students are more likely to engage in formal music training, pursue music-related degrees, or participate in advanced courses aimed at developing comprehensive musical abilities (Folkestad, 2005; Yokuş, 2020), which makes higher education an appropriate context for this review. To guide the analysis, we adopted Sternberg's (2021) concept of musical intelligence as a lens for categorizing the types of musical learning supported and musical skills developed by formative assessment.

Formative assessment

Formative assessment is often discussed in contrast to summative assessment, as both represent distinct approaches to evaluating student learning (Andrade & Brookhart, 2020; Wiliam & Thompson, 2008). Generally, summative assessment focuses on providing final grades or rankings, using standardized criteria and tests at the end of a learning period to summarize students' learning achievements. In contrast, formative assessment focuses on providing feedback to scaffold learning and enhance the learning experience through interaction between teachers and students. It uses various methods to gather and use information about student learning, emphasizing process-oriented evidence that informs teaching decisions and helps students to improve their own learning approaches (Bennett, 2011; Black & Wiliam, 1998).

Specifically, formative assessment encompasses both teacher-led and student-initiated assessment activities (self-assessment and peer assessment) throughout the educational process (Yan et al., 2021). Black and Wiliam (1998, 2009) categorized formative assessment into educational strategies initiated by three groups of individuals—teachers, peers, and learners—to drive teaching and learning. Feedback can be provided by teachers to students (i.e., teacher-directed assessment), by students to one another (i.e., peer assessment), or by students to themselves (i.e., self-assessment). When these three types of assessment are synergistically implemented, they can positively influence student learning outcomes. In practice, a large body of literature has highlighted the benefits of formative assessment across educational settings, offering strong support for its theoretical foundation. These benefits include enhanced student motivation and self-regulation (Bennett, 2011), improved learning attitudes and academic achievement (Hwang & Chang, 2011), and increased confidence and ownership of learning (Chan et al., 2014). Accordingly, formative assessment has become increasingly important in today's era of student-centered learning, as it empowers active engagement, supports educational growth, promotes power-sharing between teachers and students, and facilitates the development of self-regulated learning skills (Andrade & Heritage, 2017; Yan et al., 2021).

Formative assessment in music education

In music education, formative assessment plays a critical role in guiding musical growth (Boulet et al., 1990). First, teacher-directed assessment supports specific musical skill development by aligning instruction with learning goals; for example, Ji and O'Neill (2022) demonstrated how teachers use regular feedback to enhance students' musical practice and improve their sight-reading skills. Similarly, Hale and Green (2009) suggested using musical activities, such as sight-reading a piece or performing in an ensemble setting, as diagnostic tools, alongside rubrics to monitor musical learning progress and support the development of musicianship. Second, self-assessment fosters independence and metacognitive skills in studying music; students who reviewed their own recordings using structured criteria were better able to track tone and intonation (Goolsby, 1999), and music education students enhanced their teaching awareness through self-evaluation (Koutsoupidou, 2010). Third, peer assessment enhances collaboration and interpretive skills, as seen in Dorfman's (2021) study, where students gave feedback on technology-based performances. In addition, tools like rubrics (Thepsathit & Tangdhanakanond, 2024) and video or audio recordings (Fautley, 2013) can support both teacher- and student-led reflection. Although some performance practices are oriented toward summative evaluation (Parkes & Burrack, 2020), their formative use lies in generating actionable feedback that informs ongoing learning practices, thereby aligning with broader formative assessment principles.

However, although the potential benefits of formative assessment are increasingly recognized in music education, its implementation remains limited (Zhang et al., 2025a; Laubenthal, 2018; Valle et al., 2016). This further hinders important understandings of how formative assessment supports musical learning. Existing studies tend to focus either on the impact of formative assessment on general skills like creativity and critical thinking (Leong et al., 2012) or on specific musical skills, particularly performance (Ji & O'Neill, 2022), without examining how these domains interact. In fact, formative assessment promotes progressive learning through goal setting, reflection, and feedback. These processes primarily support students' cognitive growth beyond any specific disciplinary area (William & Thompson, 2008), giving formative assessment the potential to foster both general and domain-specific skill development across disciplines. However, how formative assessment facilitates this integration in music learning remains unclear.

Importantly, subject-specific and general skills are not separate but mutually reinforcing (Arisoy & Aybek, 2021; Sala & Gobet, 2017). For instance, developing performance skills can enhance analytical thinking, as students interpret musical structure and intent; likewise, stronger analytical or creative thinking can deepen engagement in composition or interpretation. Music education, therefore, extends beyond technical proficiency to include creativity, critical reflection, emotional expression, and social interaction. To examine how formative assessment supports this multifaceted learning, a theoretical framework is needed that bridges both general skills and subject-specific musical skills.

Theoretical foundation

Rather than applying a skill-specific theory that focuses on isolated aspects of music learning, such as performance or composition, this article adopts a broader conceptual approach. General cognitive models, such as Gardner's theory of multiple intelligences (Gardner, 2008) and Bloom's taxonomy (Bloom et al., 1956), were not considered because these frameworks lack direct alignment with the specific contexts and pedagogical practices of music education. In contrast, Sternberg (2021) and Sternberg and Kibelsbeck (2022), drawing on Gardner's (2008) generic notion of musical intelligence, offers a more integrated perspective by

identifying diverse types of musical learning that align musical skills with broader intellectual competencies.

Sternberg (2021) reframes these capacities as four interrelated skill domains: creative, analytical, practical, and wisdom-based. Although these skills are general in nature, each can be contextualized and developed through music education by aligning them with specific musical practices such as composing, performing, and critiquing. Thus, in this article, musical intelligence refers to the integrated set of general skills—creative, analytical, practical, and wisdom—that are developed and expressed through musical skill-based activities (e.g., composing a piece of music). Specifically:

- **Creative skills** enable learners to generate, invent, and innovate in music. Musical activities such as composing or playing music in novel and original ways fall under creative skills.
- **Analytical skills** support the ability to analyze, evaluate, and critique musical works. These involve enhancing musical judgment and critical thinking, such as analyzing musical structures or evaluating performances.
- **Practical skills** help students apply musical knowledge in real-world contexts such as teaching and performing. Effectively demonstrating a musical work and interacting with audiences during composition and performance also reflects practical skill development.
- **Wisdom-based skills** encourage learners to use their musical understanding in ways that serve broader ethical and societal purposes. Activities such as composing or performing music or writing critiques to promote social awareness and communal well-being exemplify wisdom-based skills.

Although these skills can be developed and assessed individually, they collectively contribute to musical intelligence, enriching learners' lives through meaningful musical experiences. This integrated model offers a comprehensive understanding of music learning and its broader impact, as illustrated in Figure 1.

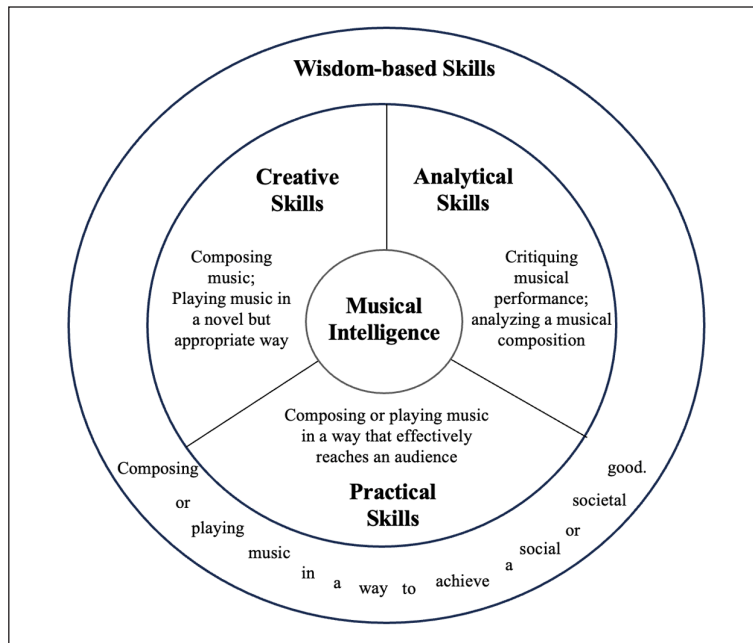
Purpose of the study

Sternberg's theory of musical intelligence can be meaningfully applied in music education; formative assessment practices—particularly those initiated by different individuals (e.g., teachers, peers, or students themselves)—may lead to varying impacts on learning outcomes. Examining single studies in isolation provides limited insight and may overlook broader patterns in how formative assessment is implemented and how it influences musical intelligence.

To date, no comprehensive review has synthesized this body of research from a comparative perspective. Therefore, this study aims to conduct a scoping review of formative assessment in music education to systematically analyze how formative assessment is applied and what effects it has on music learning, particularly in relation to the development of musical intelligence. In this article, we focus on university-level students, as they are more likely to engage in formal music training, pursue music-related degrees, or participate in advanced courses aimed at developing musical skills (Folkestad, 2005; Yokuş, 2020). As such, the use and effects of formative assessment are likely to be especially salient in this educational context.

The specific research questions guiding this review are:

RQ1. What are the characteristics of studies exploring formative assessment in higher music education?

Figure 1. A Visual Model of Sternberg's Theory of Musical Intelligence.

RQ2. How are formative assessment practices applied to music learning in higher music education?

RQ3. How do different initiator-led formative assessment activities (i.e., teacher-directed, peer-assessed, and self-assessed) impact musical intelligence in higher music education?

Methodology

To provide a holistic understanding of formative assessment in higher music education, a scoping review was conducted. Unlike systematic reviews, scoping reviews do not typically involve formal quality appraisal of included studies. Instead, it more closely resembles a literature review that systematically maps existing research on a topic to identify key concepts, research gaps, and available evidence (Mak & Thomas, 2022; Munn et al., 2018). Nevertheless, given its similarities to a systematic review, we followed Petticrew and Roberts' (2008) systematic review guidelines to ensure the trustworthiness of this study. The entire reviewing process included five steps: (1) key term selection, (2) search consultation and procedure, (3) inclusion criterion identification, (4) study quality evaluation and extraction, and (5) data analysis.

Step 1: Key term selection

Three main domains, *formative assessment*, *music education*, and *higher education*, were selected as key terms for the review process. A synonym strategy was applied to include any similar terms relating to the three domains and broaden the search range (see Table 1).

Although our synonym strategy was designed to capture a broad range of music learning contexts, it primarily focused on conventional or traditional music education terms. Terms

Table 1. Key Search Terms.

Formative assessment	Music education	Higher education
“formative assessment” OR “formative evaluation” OR “formative review” OR “assessment for learning” OR “learning-oriented assessment” OR “self-assessment” OR “peer assessment” OR “feedback”)	“music education” OR “music instruction” OR “music teaching” OR “music learning” OR “music class” OR “music lesson” OR “music tutoring” OR “music training”	“higher education” OR “college” OR “university” OR “tertiary education” OR “advanced education”

specifically related to music technology education (e.g., “music production teaching,” “digital music instruction,” “sound recording instruction”) were not included in the current search strategy. As such, studies focused on formative assessment within music technology programs may not have been captured, and we suggest that future reviews explore this increasingly relevant area of music education.

Step 2: Search strategy and procedure

The search was performed in January 2024. Before starting, the authors consulted library professionals to ensure the feasibility and effectiveness of the search strategies. The three groups of key terms were combined, and their appearance was searched in seven digital databases relevant to the education field. The initial search resulted in 893 records: PsycINFO (3), British Education Index (7), Education Research Complete (4), Scopus (373), Web of Science (3), ProQuest—Scholarly Journals (421), and ERIC (82). A further six articles were included: one article recommended by experts in the field, and five that appeared in the reference lists of the identified papers using the snowball method. After removing duplicates, 749 articles were retained and screened according to the inclusion criteria.

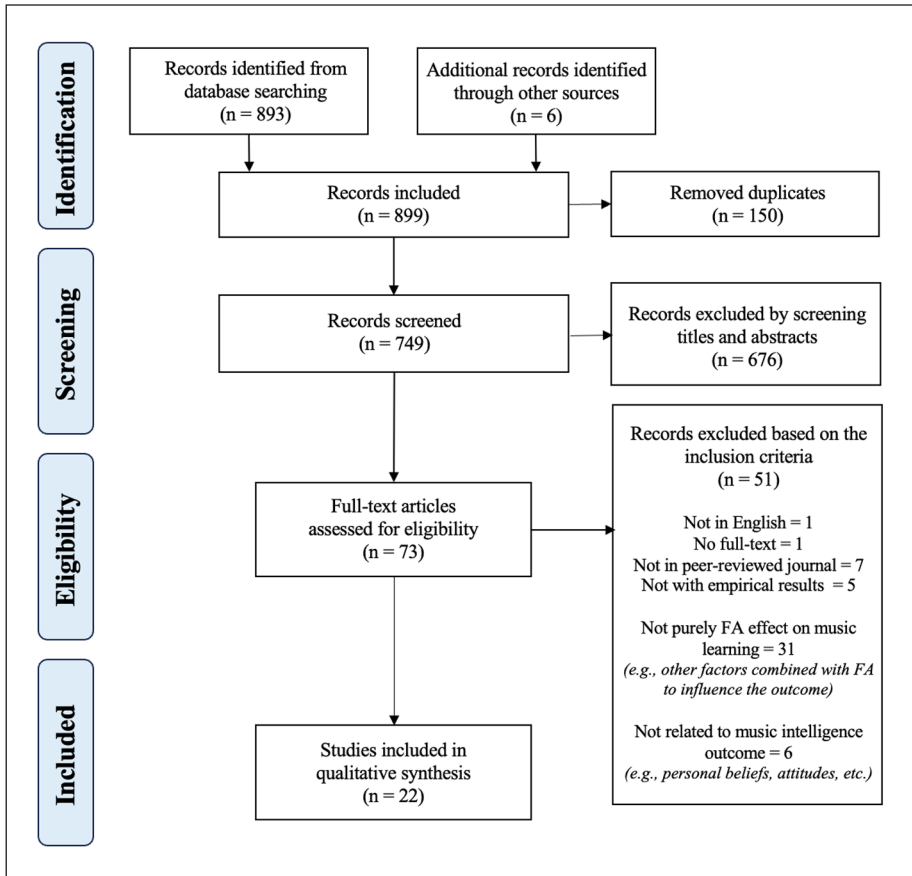
Step 3: Inclusion criteria identification

To be considered for inclusion, articles were required to satisfy the following criteria:

1. Written in English;
2. Published in peer-reviewed journals (not in books or dissertations);
3. Contained empirical results¹ (not in review or conceptual formats);
4. Involved higher education participants;
5. Reported the impact of formative assessment on music learning; and
6. Measured learning impact as the effect of formative assessment on musical intelligence (i.e., creative, analytical, practical, and wisdom-based skills). Affective and social impacts (e.g., attitudes and policy) were not considered.

Step 4: Inclusion screening, data extraction, and inter-rater reliability

To ensure the quality of the screening process, two rounds of quality screening (i.e., one for titles and abstracts, and one for full texts) were conducted. In the first round, the titles and abstracts of 20 randomly selected articles were screened by the first and third authors to reach agreement using the following options: (1) include, (2) exclude, or (3) unsure. Initial

Figure 2. Flowchart of the Processes of Literature Search and Selection.

agreement was 85%. After discussing disagreements and clarifying misunderstandings, the two raters screened another 50 randomly selected articles with 100% agreement. The subsequent screening of 749 articles resulted in 73 articles selected for full-text review.

The second round involved quality screening of the full texts. Two authors independently coded 10 randomly selected articles and reached 90% agreement on their selections. After resolving the disagreements, they screened another 20 articles and achieved 92% agreement. Following the exclusion of 51 articles that did not meet the inclusion criteria, 22 studies were retained for final analysis (see Figure 2).

After selection, the included articles were read thoroughly. A self-developed data extraction form was used to collect the following information: study characteristics (e.g., title, author, publication year, and country or region), research design (e.g., purpose, questions, and methodology), demographic details (e.g., participant characteristics and sampling methods), formative assessment strategies (e.g., teacher-led feedback, peer assessment, or self-assessment), data collection and analysis methods, findings related to musical skill development (i.e., analytical, creative, practical, and wisdom-based skills), and additional remarks (e.g., concerns, suggestions, and implications).

It is important to note that coding procedures include two coding logics. On one hand, regarding music skills coding, when coding each study's effect on musical intelligence, we only included a skill (e.g., analytical, creative, practical, or wisdom-based) if the study explicitly measured, evaluated, or reported development in that area. If a study involved activities that appeared to engage a particular skill, such as creativity or analytical analysis, but did not assess or demonstrate improvement in that domain, we did not code it as present. For example, if a study included creative tasks but did not assess students' creative skill development, we did not code it under "creative skills." Our coding thereby reflected the main focus and reported outcomes of each study, rather than the presence of activities that merely suggested potential effects. This approach ensured that the synthesis reflected the demonstrative effects of formative assessment, rather than inferred or assumed benefits. However, formative assessment strategy coding was based on whether the strategy was used as a learning activity, regardless of whether it directly led to measured outcomes. If self-assessment or peer-assessment behaviors were part of the learning process (even if not evaluated), they were coded as "included."

Overall, this data extraction form guided the systematic collection of evidence needed to address the review questions and enabled the authors to refer back consistently to the original studies during the analysis and writing phases.

Step 5: Data analysis method

The data analysis consisted of two qualitative approaches. First, deductive content analysis (Elo & Kyngäs, 2008) was applied to directly extract text related to RQ1 (i.e., general characteristics of the studies) and RQ2 (i.e., use of formative assessment practices). Second, both deductive and inductive approaches (Williams & Moser, 2019) were applied to address RQ3 (i.e., using formative assessment to develop musical intelligences). Deductive analysis was used to identify types of music intelligence based on Sternberg's (2021) categories, and inductive analysis was used to explore the specific components or factors for each type of musical intelligence (e.g., analytical intelligence for music learning). Thematic analysis (Williams & Moser, 2019) with descriptive coding in the first round and focused coding in the second round was used for this process. To ensure the reliability and consistency of the coding analysis, intercoder agreement (Belur et al., 2021) was ensured between the first and third authors. Each coder first independently coded five studies, reaching 92% agreement. The first author then coded the remaining studies alone.

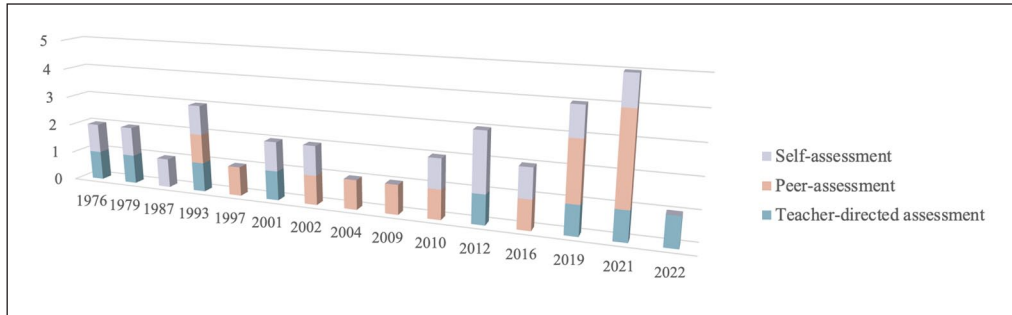
Results

RQ1: General characteristics of the studies

Overall, the findings revealed trends in general characteristics across five dimensions, which are summarized in Table S1 in the online supplementary materials. For clarity and conciseness, article numbers listed in Table S1 are used as citations when describing the results, such as Moore (1976) will be cited as Study 1.

Research timeline. The earliest documented instance of formative assessment application in teaching and learning in higher music education, featuring specific practical activities initiated by teachers, peers, and students themselves, was reported by Study 1, influenced by Yarbrough's (1974) conference manuscript. The most recent work was by Study 22.

Figure 3. Overview of Reported Use of Formative Assessment in Higher Music Education From 1976 to 2022.



Research region. The majority of the studies were conducted in the United States (9), followed by Australia (4) and the United Kingdom (3), with a few studies conducted in other countries or regions in Asia and Europe.

Research design and analysis. Qualitative research designs (13) were more prevalent than quantitative (6) and mixed-methods (3) designs. Data for quantitative analyses were commonly collected using questionnaire surveys, evaluation forms, checklists, and performance test scores, whereas interviews and students' responses were used in some qualitative studies. Recorded videos, students' portfolios, and reflective journals were also cited as data sources. Although three studies used quantitative data analysis techniques (e.g., *t*-test, correlation, or ANOVA), the majority focused on qualitative or descriptive data analysis to present their findings.

Participants. Despite the inclusion of both qualitative and quantitative research designs, the sample sizes among the studies were relatively small, ranging from six to 100 participants. The participants had diverse music backgrounds, with half of them being instrumentalists and vocalists trained to perform and others majoring in music-related fields such as music education, composition, popular music, and conducting. Although two studies included participants without a music background—one in general education (Study 1) and one in early childhood education (Study 11)—their participants were expected to teach music in the future, so they were also enrolled in university music courses and equipped with musical expertise.

RQ2: Application of formative assessment practices in higher music education

Generally, since 1976, formative assessment activities initiated by teachers (8 studies), peers (12 studies), and students themselves (12 studies) have been studied (see Figure 3). The use of peer and self-assessment in higher education was a constant topic throughout the analysed period. Compared with the majority of studies that used a single type of assessment activity (14), one study combined teacher-directed assessment with peer assessment (Study 20). Four studies applied teacher-directed assessment alongside self-assessment (Studies 1, 2, 6, 14). Two studies implemented activities involving both peer and self-assessment (Studies 7, 15). Two studies incorporated all three types of initiator-led assessments into their teaching and learning processes (Studies 4, 16).

Formative assessment practices in higher music education were examined according to four aspects (see Table S2 in the online supplementary materials): (1) setting (i.e., in what context formative assessment took place); (2) duration (i.e., for how long formative assessment practices were implemented); (3) assessment approach (i.e., how the effect of formative assessment was assessed); and (4) supporting materials (i.e., the activity-related materials involved in the process).

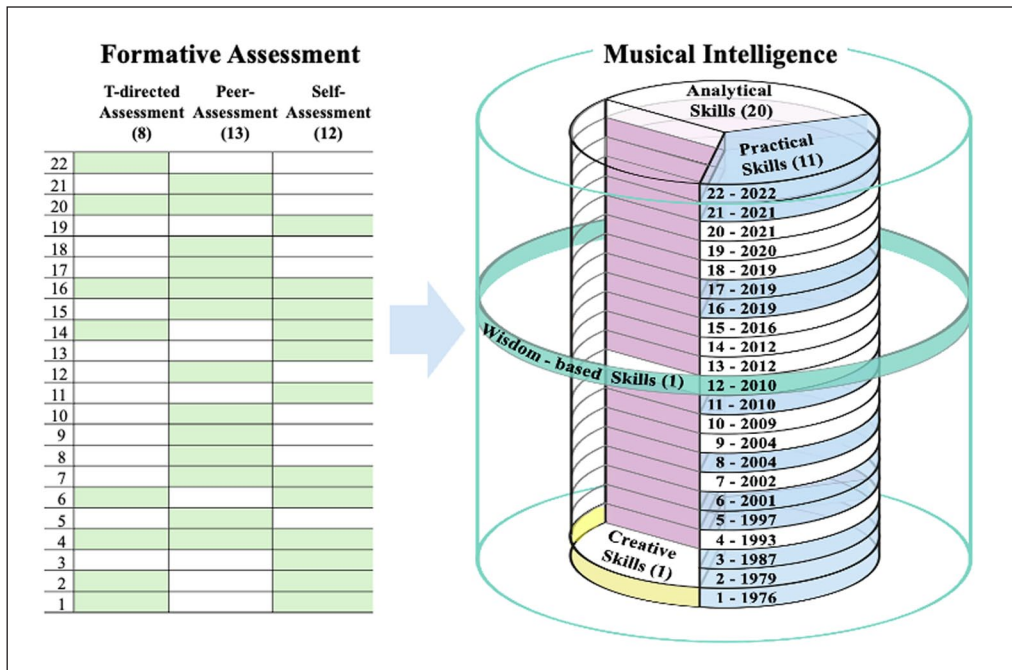
Setting. We considered both the general and the specific context of each formative assessment. First, in terms of the general context, the learning environments were either offline or online, representing traditional in-class teaching and modern technology-driven online platforms. Formative assessment activities were predominantly conducted offline before 2012. Examples included using formative assessment to observe and assess learning during courses (e.g., Studies 2, 4), evaluating performance effects (e.g., Studies 5, 8), and using formative assessment as a reflective task for critical comments and notes (e.g., Studies 7, 12). However, online formative assessment activities started to emerge from 2016 onwards. With the rapid adoption of digital technology, some in-class discussions moved from on-site to online platforms, and recorded audio and video became more accessible for evaluation in formative assessment activities. Second, in terms of the specific music learning context, the included studies primarily focused on five aspects of musical learning content: performance (12), conducting (3), music teaching (4), composition (2), and sight-reading (1). Performance, including both instrumental and vocal learning, was by far the most common study context. The review did not identify any other teaching content, such as improvisation, arrangement, analysis, or writing reviews and critiques.

Duration. Except for three studies that did not specify the duration of formative assessment implementation, all of the studies applied formative assessment over varying periods before examining its effects on learning, ranging from five days to eight years. The semester-based implementation of formative assessment was primarily due to convenience sampling strategies, in which the researchers' own students were used as participants (e.g., Studies 3, 15, 18). The eight-year implementation of formative assessment was part of a doctoral project based on a formative assessment-embedded investigation and reported findings from student interviews across four cohorts of participants (Study 10).

Assessment approach. The effects of formative assessment were assessed from three perspectives: personal perceptions (e.g., self-reported surveys, interview comments, or reflective writing), others' evaluations (e.g., teacher- or assessor-rated performance tests, observation checklists, or evaluation scores), or a combination of both. Twelve studies collected both students' perceptions and others' assessments (e.g., teacher or assessor), nine studies relied on students' personal perceptions, and one was solely based on teachers' ratings. The most common assessment method, used in almost all studies, was the use of students' personal perceptions to understand how formative assessment activities affected their learning. The only study to use teachers' ratings was the latest study (Study 22), with pre- and post-test comparisons of students' sight-reading based on instructors' assessments.

Supporting materials. When conducting formative assessment activities, nine studies used evaluation forms, tables, checklists, and rubrics as assessment tools to objectively assess learning. Not all studies used assessment criteria aligned with the evaluation process, but six studies outlined the expectations and indicators of success for the assessment. Recording equipment

Figure 4. Frequency of Use of Formative Assessment Activities and Their Impact on Musical Intelligence.



(including audio and video) was used in 14 studies and was the most frequently used tool for conducting formative assessment activities. Because the learning content in most studies was students’ musical performance, they were asked to record and watch their practice to self-reflect on their learning and achieve better self-diagnosis (e.g., Study 15) and problem identification (e.g., Study 6).

RQ3: The impact of formative assessment on students’ learning and musical intelligence

The findings indicated that initiator-led formative assessment strategies—teacher-directed, peer-assessed, and self-assessed activities—positively influenced multiple dimensions (i.e., the different aspects of musical skills) of musical intelligence (see Figure 4). Across the reviewed studies, Sternberg’s (2021) four categories of musical skills were applied as central to this impact: analytical, practical, creative, and wisdom.

In general, analytical skills were most frequently supported, with twenty of the twenty-two studies (excluding Studies 12 and 22) implementing formative assessment activities that fostered critical thinking and reflective judgment in relation to musical work and performance. Practical skills, including those related to music performance, practice, and teaching, were enhanced in eleven studies (Studies 1, 2, 3, 5, 6, 8, 11, 16, 17, 21, 22). Creative development was reported in one study (Study 1), which combined all three formative assessment strategies. In addition, peer assessment in Study 12 contributed to students’ ethical awareness and sense of social responsibility.

To further explore the relationship between formative assessment and students’ music learning experience, the following sections report (1) students’ perceived benefits of

participating in formative assessment activities, (2) the specific mechanisms through which formative assessment supports the development of musical skills, and (3) the remaining concerns and difficulties of applying formative assessment strategies.

Students' perceived benefits of formative assessment activities. In the reviewed studies, students were active participants in formative assessment activities within higher music education. Their perceptions, whether self-reported or observed by instructors, offered insights into the pedagogical value and effectiveness of formative assessment. Thematic analysis of the data revealed three major areas in which students experienced positive affect through teacher-directed, peer-assessed, and self-assessed formative assessment practices.

First, students widely recognized *formative assessment as a tool for improving specific musical skills and overall performance*. Teacher-directed strategies were particularly valued for providing clear structure, targeted feedback, and reduced performance anxiety, which helped students focus on skill development and preparation (Studies 1, 2, 3, 6, 21, 22). Similarly, self-assessment practices, such as video reflection and technical checklists, were seen as essential for enhancing performance accuracy and supporting independent practice (Studies 4, 6, 14, 15, 19).

Second, students reported that *formative assessment activities significantly contributed to the development of critical thinking and evaluative skills*. Peer assessment, in particular, encouraged deeper understanding of musical criteria and integration of theory and practice. Engaging in structured feedback processes helped students refine their judgment and apply analytical reasoning to both their own and others' performances (Studies 5, 8, 9, 10, 12, 17, 20). Self-assessment also supported this process by prompting reflection and comparison with professional standards (Studies 12, 15, 21).

Third, *formative assessment was perceived as instrumental in fostering motivation, confidence, and a sense of ownership over learning*. Students described how giving and receiving peer feedback increased their engagement and responsibility, especially when they saw their suggestions contribute to real improvement (Studies 5, 9, 10, 12, 17, 18). Self-assessment similarly promoted independence and self-regulation, with students reporting greater self-awareness and metacognitive control over their learning processes (Studies 2, 6, 7, 13, 14). Additional studies highlighted the role of self-assessment in building confidence and reducing anxiety, particularly in real-world teaching or performance settings (Studies 11, 18), and in fostering a stronger sense of ownership (Studies 6, 7, 11, 16).

The impact of formative assessment strategies on musical intelligence. Generally, all three initiator-led formative assessment strategies—teacher-directed, peer, and self-assessment—had positive effects on the four components of musical intelligence. Table 2 summarizes these impacts and highlights newly identified formats of formative assessment that further support the development of musical skills.

Each strategy activated specific mechanisms of action that shaped the instructional process and supported the development of musical intelligence. The pedagogical functions of formative assessment—how it was practically applied to guide learning, improve performance, and cultivate evaluative thinking in the context of music instruction—could be specifically summarized in three aspects.

First, teacher-directed assessment primarily supported learning through modeling, guided critique, and targeted feedback. Instructors facilitated analytical development by modeling teaching or musical behaviors and leading structured reflection or discussion (Studies 1, 2, 4, 6, 10, 14, 16). Teacher comments during or after student performances helped learners identify technical issues such as gesture, timing, or coordination (Studies 2, 4, 6, 14, 16). This

Table 2. Detailed Impact of Using Formative Assessment Activities on Musical Intelligence.

Formative assessment strategy	Mechanism of action in formative assessment	Skills	Included studies
Teacher-directed	Modeling, guided discussions, structured feedback on philosophical and technical topics	Analytical	1, 2, 4, 6, 10, 14, 16
	Feedback on conducting, rhythm, teaching effectiveness, and music delivery	Practical	1, 2, 6, 16, 22
	Encouraging creative teaching strategies via rubric-based evaluation	Creative	1
Peer assessment	Structured peer critique using rubrics, observations, and feedback discussions	Analytical	4, 7, 8, 15, 16, 17, 18, 20, 21
	Performance-based peer evaluation in live or practice-based settings	Practical	5, 8, 16, 17, 21
	Ethical dialogue in marking peers, fairness debates, professional responsibility	Wisdom	12
Self-assessment	Structured self-observation, reflection via checklists, video, or written critique	Analytical	1, 2, 3, 4, 6, 7, 13, 14, 15, 16, 19
	Evaluation of own teaching or performance in real-world contexts	Practical	1, 2, 3, 6, 11, 16

feedback was then applied in authentic teaching or performance contexts, including live conducting, classroom instruction, and instrumental playing, supporting practical skill development (Studies 1, 2, 6, 16, 22). In one case, creativity was addressed explicitly through rubric-based evaluation of students' original teaching strategies (Study 1).

Peer assessment was implemented through structured evaluation tasks and, in one case, ethical dialogue. Students used rubrics or observation tools to assess peer performances or teaching demonstrations, developing musical-analytical skills such as diagnosing technique and expression (Studies 4, 7, 8, 15, 16, 17, 18, 20, 21). In some studies, peer assessment took place within authentic or applied contexts, such as concerts, rehearsals, or teaching practice, enhancing its relevance to practical music-making (Studies 5, 8, 16, 17, 21). Ethical reflection was uniquely emphasized in one study, where students discussed fairness, bias, and professional responsibility in peer evaluation, contributing to wisdom-based learning (Study 12).

Finally, self-assessed formative evaluation was applied through video-based observation, written critique, and structured self-monitoring, consistently fostering analytical skill development. Students analyzed their own teaching or performance using observation forms, rubrics, or reflective writing, helping them identify strengths, weaknesses, and areas for improvement (Studies 1, 2, 3, 4, 6, 7, 13, 14, 15, 16, 19). In several cases, these reflections were tied to real-world tasks such as public teaching or live performance, supporting practical application of musical knowledge (Studies 1, 2, 3, 6, 11, 16).

Remaining concerns and difficulties in the formative assessment application. Although formative assessment has demonstrated a positive impact on musical learning, several studies also highlighted student concerns and practical challenges related to its implementation. These concerns suggested that receptivity to the different formative assessment strategies was not simply

universal and that their effectiveness depended on contextual and pedagogical support. The primary challenge was the time-consuming nature of self- and peer assessment. Students often felt overwhelmed when evaluating multiple aspects of their musical performance (Study 3), and peer assessment required them to invest time in becoming familiar with the assessment criteria and developing confidence in using a professional musician's perspective to evaluate their peers (Study 5).

Another concern was the need for professional guidance to support students as assessors of musical performance (Studies 12, 17). Without appropriate guidance, students may misunderstand the required music knowledge, leading to biased, self-critical analysis and difficulties in breaking down assessment criteria into specific musical components. There is, therefore, a risk that students may rely on personal perceptions rather than objective professional criteria when evaluating the musical work of their peers, potentially compromising the validity of the assessment results.

A final noteworthy concern was that self-critique required students to be highly sensitive in identifying facts and conducting objective analyses, potentially impacting their self-esteem. Study 6 questioned "whether self-critical analysis for some may in fact counteract creativity" (p. 224). This underscores the importance of closely monitoring the assessment process and ensuring that students' creative abilities remain unaffected by their critical self-analysis.

Discussion

An overview of applying formative assessment in higher music education

This scoping review provides an overview of formative assessment studies conducted in higher music education contexts and examines the impact of formative assessment on college students' development of musical skills. Since 1976, assessment activities initiated by teachers, peers, and students have been explored as three initiator-led strategies within higher music learning contexts. Five out of 22 studies investigated the effects of a single strategy (teacher-led, peer, or self-assessment), while 17 studies examined combined strategies. All studies reported their positive impacts on students' musical development.

From students' perspectives, their positive attitudes suggested that applying formative assessment strategies enhanced their musical learning outcomes and performance, supported the development of critical thinking and self-reflection, and boosted confidence and motivation. These findings, from a music learning perspective, are consistent with prior research highlighting formative assessment's role in improving student motivation and self-regulation (Bennett, 2011), learning attitudes and achievement (Hwang & Chang, 2011), and confidence and ownership of learning (Chan et al., 2014). This trend also aligns with Black and Wiliam's (1998, 2009) view that diverse formative assessment strategies can positively influence learning, with greater impact when used synergistically.

From the perspectives of researchers of the included studies, different formative assessment strategies supported various aspects of musical skill development, particularly in enhancing musical thinking, practice, and performance. These effects could be discussed through two main lenses: self-analysis and problem analysis. Self-analysis illustrates how formative assessment fostered metacognitive awareness. Learners engaged in self-reflection, critique, and correction, leading to greater self-understanding and more intentional practice (Yan, 2022). When guided by clear evaluative criteria, students developed autonomy and emotional engagement in music learning (Boucher et al., 2021), supporting earlier findings on formative assessment's role in helping learners monitor their progress and development (Goolsby, 1999;

Koutsoupidou, 2010). Problem analysis focuses on how formative assessment helped diagnose and address specific performance issues. In the reviewed studies, peer feedback encouraged critical listening and the identification of technical or behavioral problems (Daniel, 2001; Dorfman, 2021). Teacher feedback supported students in refining skills such as performance and sight-reading (Hale & Green, 2009; Ji & O'Neill, 2022), whereas video recording allowed learners to compare and evaluate their performances more critically, enhancing their problem-solving skills (Fautley, 2013; Liu et al., 2021). These practices reflected formative assessment's process-oriented nature, where assessment is used not only to evaluate but also to reflect gaps in understanding and to inform and improve learning (Andrade & Heritage, 2017; Bennett, 2011).

The specific impact of formative assessment on musical intelligence

In applying formative assessment strategies to support the four skill-based aspects of musical intelligence—analytical skills, practical skills, creative skills, and wisdom—results indicated that all three initiator-led formative assessment strategies positively impacted skill development. Teacher-directed assessment and self-assessment primarily supported analytical and practical skills, whereas peer assessment contributed across all four dimensions. This reinforces formative assessment's effectiveness as an in-class learning tool for diverse educational contexts, including music education (Black & Wiliam, 1998). According to Sternberg (2021), these four aspects interact to foster well-rounded musicianship, and evidence suggests that formative assessment can comprehensively support these aspects, underscoring its value in music learning.

However, formative assessment's uneven impact across the four skills merits attention. It was most frequently reported as enhancing analytical and practical skills. This emphasis may reflect formative assessment's inherent focus on stimulating reflection and self-regulation (Wiliam & Thompson, 2008). Similar trends have appeared in other fields: peer assessment has been reported to promote critical feedback in English writing tasks (Lu & Bol, 2007), and self-assessment deepened analytical thinking in teacher education (Yan & Brown, 2017). Thus, formative assessment's role in fostering critical thinking seems consistent across disciplines. In addition, over half of the reviewed studies applied formative assessment to practical skills, with improvements reported in musical performance (Liu et al., 2021), teaching (Legette & Royo, 2021), and composition (Searby & Ewers, 1997). This aligns with wider findings that formative assessment enhances academic performance (Yan & Chiu, 2023). In music education, activities like performance, composition, and teaching demonstration parallel academic tasks, making formative assessment a natural strategy for practical skill development.

By contrast, few studies addressed formative assessment's role in promoting creative skills and wisdom. Only peer assessment was reported to encourage creativity (Moore, 1976) and ethical dialogue (Latukeyu, 2010); teacher-directed and self-assessment were not reported to support these areas. This raises a critical question: Is formative assessment inherently limited in fostering creativity and wisdom, or could its formats be adapted?

Evidence from other disciplines shows that formative assessment can support creativity. Creativity-based rubrics have improved creative skills in mathematics (Savic et al., 2017), and mind-mapping has enhanced creative self-efficacy in language learning (Yan et al., 2022). As creativity is central to music learning (Kokotsaki & Hallam, 2007; Parkes & Burrack, 2020), conventional formative assessment formats—mainly oral or written feedback—may not align with music's exploratory, process-based nature. Although musical creativity is often evaluated in summative assessment, it is embedded in performance, interpretation, and composition

(Burnard, 2012), which could be better supported by creative formative assessment tools like mind maps (Zhang & Leung., 2023). In addition, digital platforms now enable real-time, collaborative feedback (Yoo, 2016), and Artificial Intelligence (AI) tools offer new ways to track and support students' creative processes (L. Chen, 2024). These technologies make it more feasible to apply formative assessment throughout the creative journey, not just at the end. Thus, the limited attention to creativity in the reviewed studies may reflect the constraints of past research and search terms rather than formative assessment's actual capacity. We encourage further exploration of how formative assessment can be innovatively applied to foster musical creativity.

Regarding wisdom, few existing studies explored how assessment cultivates broader ethical or societal perspectives. Yet, educational goals increasingly emphasize broad and general competencies such as collaboration, critical thinking, and problem-solving (UNESCO, 2015). This shift highlights the need for scaffolding formative practices that help students connect learning to broader societal values. As formative assessment has shown potential in promoting ethical and responsible thinking (Latukeyu, 2010), it should be more consciously applied in music education to develop socially responsible musicians.

Challenges and opportunities for applying formative assessment in higher music education

This review suggests that formative assessment remains underused in higher music education compared with general education. For example, Morris et al. (2021) identified 28 studies on formative assessment in higher education between 2000 and 2019, while this review, which focused exclusively on higher music education and covered a broader time span (1976–2023), identified only 22 relevant studies. This gap raises questions about the limited uptake or undervaluing of formative assessment in music conservatoires and university music programs. Several challenges may explain this underutilization.

First, time constraints are a major challenge, as formative assessment requires continuous feedback, reflection, and instructional adjustment (Andrade & Brookhart, 2020). In higher education settings, where contact hours are limited and lectures often prioritize content delivery over iterative skill development (Altbach & Knight, 2007; Naseer et al., 2024), it can be difficult for teachers to facilitate peer and self-assessment in large or performance-focused classes. However, these barriers may be mitigated by embedding formative assessment within existing teaching structures, such as studio lessons, rehearsals, and masterclasses, through brief peer feedback, guided reflection, and formative rubrics, without significantly adding to instructional time. Digital tools like video-based feedback (Boucher et al., 2021) or online platforms (Yoo, 2016) may help streamline this process.

A second challenge concerns students' readiness to serve as assessors of their own or others' work. Rooted in conservatoire traditions and Western classical ideals, music education often emphasizes expert judgment and high-performance standards (Parkes & Burrack, 2020). This can lead to skepticism about the validity of peer and self-assessment, particularly in subjective domains such as musical interpretation. Bennett (2011) describes this as a "domain dependency" issue, where formative assessment may be more readily applied in fields with objective criteria. Yet, evidence shows that when students are explicitly taught how to assess using clear criteria, and when reflective practices are scaffolded over time, their confidence and evaluative skills improve (Yan, 2022). Introducing assessment literacy into early coursework could help

normalize these roles and encourage a more collaborative learning environment that challenges the traditional expert–novice divide.

Third, self-assessment in music might raise psychological concerns. Some students may experience anxiety or diminished creative confidence when engaging in critical self-reflection (Daniel, 2001; Sun, 2022). Meanwhile, creativity in music thrives when both exploration and structured reflection are supported (Burnard, 2012). These differing perspectives suggest that formative assessment must be implemented in ways that support, rather than constrain, artistic growth. Teachers play a crucial role in creating environments that strike a balance between critical thinking and creative freedom.

Ultimately, many of these challenges point to a broader issue: Are music educators adequately prepared to adopt formative assessment in their teaching? Prior studies indicate that teacher training in formative assessment remains limited, especially in performance-based programs where pedagogical preparation is not always emphasized (Russell & Austin, 2010; Zhang et al., 2025a, 2025b). Without appropriate training, teachers may lack the confidence or strategies to integrate formative assessment into their practice. Addressing this gap requires targeted professional development that models effective formative assessment techniques and aligns with authentic assessment approaches. These include portfolios, collaborative projects, and iterative feedback, which mirror real-world music-making and support students' professional and lifelong learning (Andrade & Brookhart, 2020). Institutional support is also essential, as recognizing and rewarding reflective and innovative teaching can help shift assessment cultures and make formative assessment a more sustainable part of music education.

Conclusion

This scoping review systematically evaluated the application and impact of formative assessment in higher music education by analyzing 22 empirical studies conducted since the 1970s. The results indicate that three types of formative assessment activities (i.e., teacher-led, peer, and self-assessments) have been implemented in this field, either individually or in combination, and have contributed to the development of musical intelligence. The positive influence of formative assessment is particularly evident in its ability to enhance students' analytical and practical music skills, fostering deeper thinking and the practical application of knowledge. However, this highlights a limited focus on the use of formative assessment to nurture music students' creativity and wisdom.

Although this study aimed to provide a comprehensive overview of formative assessment's influence on musical intelligence, some limitations remain. For instance, the terminology used to describe formative assessment in music education has evolved over time, and emerging terms such as digital assessment or AI-based feedback systems may have been overlooked. In addition, as assessment strategies are often closely tied to local contexts, the exclusion of non-English studies may limit the review's representation of formative assessment practices in different regions.

Future research could explore a wider range of formative assessment strategies, particularly those enhanced by technology, to examine whether formative assessment is evolving in current music education contexts. Qualitative studies may offer valuable insights into how university and school music teachers perceive and implement formative assessment to support musical learning. Intervention studies could be designed to develop targeted materials that foster students' creative skills and wisdom, and—although this review found no clear differences in the implementation of formative assessment across regions—comparative studies could provide deeper understanding of its effectiveness across diverse educational settings.

Author contributions

Le-Xuan Zhang: Conceptualization; Data curation; Formal analysis; Funding acquisition; Investigation; Methodology; Project administration; Visualization; Writing—original draft.

Zi Yan: Conceptualization; Methodology; Resources; Supervision; Validation; Writing—review & editing.

Xiang Wang: Formal analysis; Investigation; Resources; Validation; Writing—review & editing.

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research study was supported by the Start-Up Research Grant (number RG 19/2023-2024R) from The Education University of Hong Kong.

ORCID iDs

Le-Xuan Zhang  <https://orcid.org/0000-0002-4394-7734>

Zi Yan  <https://orcid.org/0000-0001-9305-884X>

Supplemental material

Supplemental material for this article is available online.

Note

1. In this review, “empirical” refers to studies that involved systematic data collection and analysis to answer research questions. We excluded theoretical papers, literature reviews, and book chapters (including those from edited volumes such as the *Oxford Handbooks*), as they did not typically present original empirical findings.

References

References marked with an asterisk () indicate articles included in the review analysis.*

- Altbach, P., & Knight, J. (2007). The internationalization of higher education: Motivations and realities. *Journal of Studies in International Education*, 11(3–4), 290–305. <https://doi.org/10.1177/1028315307303542>
- Andrade, H. L., & Brookhart, S. M. (2020). Classroom assessment as the co-regulation of learning. *Assessment in Education: Principles, Policy & Practice*, 27(4), 350–372. <https://doi.org/10.1080/0969594X.2019.1571992>
- Andrade, H. L., & Heritage, M. (2017). *Using formative assessment to enhance learning, achievement, and academic self-regulation*. Routledge.
- Ariso, B., & Aybek, B. (2021). The effects of subject-based critical thinking education in mathematics on students' critical thinking skills and virtues. *Eurasian Journal of Educational Research*, 92, 99–119. <https://doi.org/10.14689/ejer.2021.92.6>
- Belur, J., Tompson, L., Thornton, A., & Simon, M. (2021). Interrater reliability in systematic review methodology: exploring variation in coder decision-making. *Sociological Methods & Research*, 50(2), 837–865. <https://doi.org/10.1177/0049124118799372>
- Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice*, 18(1), 5–25. <https://doi.org/10.1080/0969594X.2010.513678>

- *Bergee, M. J., & Cecconi-Roberts, L. (2002). Effects of small-group peer interaction on self-evaluation of music performance. *Journal of Research in Music Education*, 50(3), 256–268. <https://doi.org/10.2307/3345802>
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7–74. <https://doi.org/10.1080/0969595980050102>
- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21(1), 5–31. <https://doi.org/10.1007/s11092-008-9068-5>
- *Blom, D., & Poole, K. (2004). Peer assessment of tertiary music performance: Opportunities for understanding performance assessment and performing through experience and self-reflection. *British Journal of Music Education*, 21(1), 111–125. <https://doi.org/10.1017/S0265051703005539>
- Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *Taxonomy of educational objectives: The classification of educational goals: Handbook 1: Cognitive domain*. New York: Longmans.
- *Boucher, M., Creech, A., & Dubé, F. (2021). Video feedback and the self-evaluation of college-level guitarists during individual practice. *Psychology of Music*, 49(2), 159–176. <https://doi.org/10.1177/0305735619842374>
- Boulet, M. M., Simard, G., & De Melo, D. (1990). Formative evaluation effects on learning music. *The Journal of Educational Research*, 84(2), 119–125. <https://doi.org/10.1080/00220671.1990.10886002>
- Burnard, P. (2012). *Musical creativities in practice*. Oxford University Press.
- Chan, P. E., Graham-Day, K. J., Ressa, V. A., Peters, M. T., & Konrad, M. (2014). Beyond involvement: Promoting student ownership of learning in classrooms. *Intervention in School and Clinic*, 50(2), 105–113. <https://doi.org/10.1177/1053451214536039>
- Chen, L. (2024). Unlocking the beat: How AI tools drive music students' motivation, engagement, creativity and learning success. *European Journal of Education*, 60(1), e12823. <https://doi.org/10.1111/ejed.12823>
- Chen, P. P., & Bonner, S. M. (2019). A framework for classroom assessment, learning, and self-regulation. *Assessment in Education: Principles, Policy & Practice*, 27(4), 373–393. <https://doi.org/10.1080/0969594X.2019.1619515>
- Conway, C. M. (2020). *Teaching music in higher education*. Oxford University Press.
- *Cremaschi, A. M. (2012). The effect of a practice checklist on practice strategies, practice self-regulation and achievement of collegiate music majors enrolled in a beginning class piano course. *Research Studies in Music Education*, 34(2), 223–233. <https://doi.org/10.1177/1321103X12464743>
- *Daniel, R. (2001). Self-assessment in performance. *British Journal of Music Education*, 18(3), 215–226. <https://doi.org/10.1017/S0265051701000316>
- *Daniel, R. (2004). Peer assessment in musical performance: The development, trial and evaluation of a methodology for the Australian tertiary environment. *British Journal of Music Education*, 21(1), 89–110. <https://doi.org/10.1017/S0265051703005515>
- *Deniz, J. (2012). Video recorded feedback for self-regulation of prospective music teachers in piano lessons. *Journal of Instructional Psychology*, 39(1), 17–25. <https://www.proquest.com/scholarly-journals/video-recorded-feedback-self-regulation/docview/1081191358/se-2>
- *Dorfman, J. (2021). Project peer assessment in a preservice music technology course. *Contributions to Music Education*, 46, 71–90. <https://www.jstor.org/stable/27125931>
- Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *Journal of Advanced Nursing*, 62(1), 107–115. <https://doi.org/10.1111/j.1365-2648.2007.04569.x>
- Fautley, M. (2013). The potential of audio and video for formative assessment purposes in music education in the lower secondary school in England: Issues arising from a small-scale study of trainee music teachers. *Journal of Music, Technology and Education*, 6(1), 29–42. https://doi.org/10.1386/jmte.6.1.29_1
- Folkestad, G. (2005). Here, there and everywhere: Music education research in a globalised world. *Music Education Research*, 7(3), 279–287. <https://doi.org/10.1080/14613800500324390>
- Gardner, H. E. (2008). *Multiple intelligences: New horizons in theory and practice*. Basic books.
- Goolsby, T. W. (1999). Assessment in Instrumental Music: How can band, orchestra, and instrumental ensemble directors best assess their student's learning? Here are some evaluation tools and techniques to consider. *Music Educators Journal*, 86(2), 31–50. <https://doi.org/10.2307/3399587>

- Hale, C. L., & Green, S. K. (2009). Six key principles for music assessment. *Music Educators Journal*, 95(4), 27–31. <https://doi.org/10.1177/0027432109334772>
- *Hwang, G. H., Chen, B., & Sung, C. W. (2019). Impacts of flipped classrooms with peer assessment on students' effectiveness of playing musical instruments—taking amateur erhu learners as an example. *Interactive Learning Environments*, 27(8), 1047–1061. <https://doi.org/10.1080/10494820.2018.1481105>
- Hwang, G. J., & Chang, H. F. (2011). A formative assessment-based mobile learning approach to improving the learning attitudes and achievements of students. *Computers & Education*, 56(4), 1023–1031. <https://doi.org/10.1016/j.compedu.2010.12.002>
- *Ji, W., & O'Neill, K. (2022). Effectiveness of targeted feedback in improving rhythm sightreading. *Music Education Research*, 24(4), 482–493. <https://doi.org/10.1080/14613808.2022.2069233>
- *Johnston, H. (1993). The use of video self-assessment, peer-assessment, and instructor feedback in evaluating conducting skills in music student teachers. *British Journal of Music Education*, 10(1), 57–63. <https://doi.org/10.1017/S0265051700001431>
- Kokotsaki, D., & Hallam, S. (2007). Higher education music students' perceptions of the benefits of participative music making. *Music Education Research*, 9(1), 93–109. <https://doi.org/10.1080/14613800601127577>
- *Koutsoupidou, T. (2010). Self-assessment in generalist preservice kindergarten teachers' education: Insights on training, ability, environments, and policies. *Arts Education Policy Review*, 111(3), 105–111. <https://doi.org/10.1080/10632911003626937>
- *Latukefu, L. (2010). Peer assessment in tertiary level singing: Changing and shaping culture through social interaction. *Research Studies in Music Education*, 32(1), 61–73. <https://doi.org/10.1177/1321103X10370091>
- Laubenthal, J. (2018). Assessment in the private studio setting: Supporting student learning, providing effective instruction, and building faculty-student interaction. *Music Educators Journal*, 104(3), 54–59. <https://doi.org/10.1177/0027432117745139>
- *Legette, R. M., & Royo, J. (2021). Pre-service music teacher perceptions of peer feedback. *Research Studies in Music Education*, 43(1), 22–38. <https://doi.org/10.1177/1321103X19862298>
- Leong, S., Burnard, P., Jeanneret, N., Leung, B. W., & Waugh, C. (2012). Assessing creativity in music: International perspectives and practices. In G. E. McPherson & G. F. Welch (Eds.), *Creativities, technologies, and media in music learning and teaching: An Oxford handbook of music education* (Vol. 5, pp. 389–407). Oxford University Press.
- *Liu, C., Wan, P., Tu, Y. F., Chen, K., & Wang, Y. (2021). A WSQ-based mobile peer assessment approach to enhancing university students' vocal music skills and learning perceptions. *Australasian Journal of Educational Technology*, 37(6), 1–17. <https://doi.org/10.14742/ajet.6832>
- Lu, R., & Bol, L. (2007). A comparison of anonymous versus identifiable e-peer review on college student writing performance and the extent of critical feedback. *Journal of Interactive Online Learning*, 6(2), 100–115. https://digitalcommons.odu.edu/cgi/viewcontent.cgi?article=1002&context=efl_fac_pubs
- Mak, S., & Thomas, A. (2022). Steps for conducting a scoping review. *Journal of Graduate Medical Education*, 14(5), 565–567. <https://doi.org/10.4300/JGME-D-22-00621.1>
- *Moore, R. S. (1976). Effect of differential teaching techniques on achievement, attitude, and teaching skills. *Journal of Research in Music Education*, 24(3), 129–141. <https://doi.org/10.2307/3345156>
- Morris, R., Perry, T., & Wardle, L. (2021). Formative assessment and feedback for learning in higher education: A systematic review. *Review of Education*, 9(3), e3292. <https://doi.org/10.1002/rev3.3292>
- Munn, Z., Peters, M. D., Stern, C., Tufanaru, C., McArthur, A., & Aromataris, E. (2018). Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. *BMC Medical Research Methodology*, 18, 1–7. <https://doi.org/10.1186/s12874-018-0611-x>

- Naseer, F., Khan, M., Tahir, M., Addas, A., & Aejaaz, S. (2024). Integrating deep learning techniques for personalized learning pathways in higher education. *Heliyon*, *10*(11), e32628. <https://doi.org/10.1016/j.heliyon.2024.e32628>
- Parkes, K. A., & Burrack, F. (2020). *Developing and applying assessments in the music classroom*. Routledge.
- Petticrew, M., & Roberts, H. (2008). *Systematic reviews in the social sciences: A practical guide*. John Wiley.
- *Pulman, M. (2009). Seeing yourself as others see you: Developing personal attributes in the group rehearsal. *British Journal of Music Education*, *26*(2), 117–135. <https://doi.org/10.1017/S0265051709008390>
- Russell, J. A., & Austin, J. R. (2010). Assessment practices of secondary music teachers. *Journal of Research in Music Education*, *58*(1), 37–54. <https://doi.org/10.1177/0022429409360062>
- Sala, G., & Gobet, F. (2017). When the music's over. Does music skill transfer to children's and young adolescents' cognitive and academic skills? A meta-analysis. *Educational Research Review*, *20*, 55–67. <https://doi.org/10.1016/j.edurev.2016.11.005>
- Savic, M., Karakok, G., Tang, G., El Turkey, H., & Naccarato, E. (2017). Formative assessment of creativity in undergraduate mathematics: Using a creativity-in-progress rubric (CPR) on proving. In R. Leikin & B. Sriraman (Eds.), *Creativity and giftedness: Interdisciplinary perspectives from mathematics and beyond* (pp. 23–46). Springer Nature.
- *Searby, M., & Ewers, T. (1997). An evaluation of the use of peer assessment in higher education: A case study in the School of Music, Kingston University. *Assessment & Evaluation in Higher Education*, *22*(4), 371–383. <https://doi.org/10.1080/0260293970220402>
- Sternberg, R. J. (2021). Toward a theory of musical intelligence. *Psychology of Music*, *49*(6), 1775–1785. <https://doi.org/10.1177/0305735620963765>
- Sternberg, R. J., & Kibelsbeck, J. G. (2022). Teaching musical learning as problem-solving: Applying a theory of musical intelligence to musical instruction. *Psychology of Music*, *50*(5), 1696–1706. <https://doi.org/10.1177/03057356211055215>
- Sun, J. (2022). Exploring the impact of music education on the psychological and academic outcomes of students: Mediating role of self-efficacy and self-esteem. *Frontiers in Psychology*, *13*, 1–13. <https://doi.org/10.3389/fpsyg.2022.841204>
- Thepsathit, P., & Tangdhanakanond, K. (2024). The development of formative assessment rubrics for enhancing students' performance on Thai percussion instruments. *International Journal of Music Education*, *42*(4), 674–690. <https://doi.org/10.1177/02557614231192189>
- UNESCO. (2015). *Education 2030 Framework for Action*. <https://apa.sdg4education2030.org/education-2030-framework-action>
- Valle, C., Andrade, H., Palma, M., & Hefferen, J. (2016). Applications of peer assessment and self-assessment in music. *Music Educators Journal*, *102*(4), 41–49. <https://doi.org/10.1177/0027432116644652>
- West, C., & Clauhs, M. (2019). Peers observing peers: Exploring cross-institutional observations and dialogues facilitated by videoconferencing software. *Journal of Music Teacher Education*, *28*(2), 43–55. <https://doi.org/10.1177/1057083718803633>
- William, D., & Thompson, M. (2008). Integrating assessment with instruction: What will it take to make it work? In C. A. Dwyer (Ed.), *The future of assessment: Shaping teaching and learning* (pp. 53–82). Erlbaum.
- Williams, M., & Moser, T. (2019). The art of coding and thematic exploration in qualitative research. *International Management Review*, *15*(1), 45–55. <https://imrjournal.org/uploads/1/4/2/8/14286482/imr-v15n1art4.pdf>
- Yan, Z. (2022). *Student self-assessment as a process for learning*. Routledge.
- Yan, Z., & Brown, G. T. (2017). A cyclical self-assessment process: Towards a model of how students engage in self-assessment. *Assessment & Evaluation in Higher Education*, *42*(8), 1247–1262. <https://doi.org/10.1080/02602938.2016.1260091>
- Yan, Z., & Chiu, M. M. (2023). The relationship between formative assessment and reading achievement: A multilevel analysis of students in 19 countries/regions. *British Educational Research Journal*, *49*(1), 186–208. <https://doi.org/10.1002/berj.3837>

- Yan, Z., Lee, J. C. K., Hui, S. K. F., & Lao, H. (2022). Enhancing students' self-efficacy in creativity and learning performance in the context of English learning: The use of self-assessment mind maps. *Frontiers in Psychology*, 13, 871781. <https://doi.org/10.3389/fpsyg.2022.871781>
- Yan, Z., Li, Z., Panadero, E., Yang, M., Yang, L., & Lao, H. (2021). A systematic review on factors influencing teachers' intentions and implementations regarding formative assessment. *Assessment in Education: Principles, Policy & Practice*, 28(3), 228–260. <https://doi.org/10.1080/0969594X.2021.1884042>
- Yarbrough, C. (1976, March). *The effect of videotaped observation and self-evaluation on rehearsal behavior of student conductors* [Conference presentation]. National Biennial In-Service Conference of the Music Educators National Conference, Atlantic City, NJ.
- *Yarbrough, C. M. (1987). The relationship of behavioral self-assessment to the achievement of basic conducting skills. *Journal of Research in Music Education*, 35(3), 183–189. <https://doi.org/10.2307/3344960>
- *Yarbrough, C. M., Wapnick, J., & Kelly, R. (1979). Effect of videotape feedback techniques on performance, verbalization, and attitude of beginning conductors. *Journal of Research in Music Education*, 27(2), 103–112. <https://doi.org/10.2307/3344896>
- Yokuş, T. (2020). The effect of metacognitive strategies-based teaching practice in guitar education on performance achievement. *Psychology of Music*, 49(6), 1605–1619. <https://doi.org/10.1177/0305735620968259>
- *Yoo, H. (2016). A web-based environment for facilitating reflective self-assessment of choral conducting students. *Contributions to Music Education*, 41, 113–130. <https://www.jstor.org/stable/24711131>
- Zhang, L. X., & Leung, B. W. (2023). Context matters: Adaptation of student-centred education in China school music classrooms. *Music Education Research*, 25(4), 418–434. <https://doi.org/10.1080/14613808.2023.2230587>
- Zhang, L. X., Yan, Z., & Wang, X. (2025a). Understanding music teachers' formative assessment intention and implementation: A Chinese Mainland context. *International Journal of Music Education Advance* online publication. <https://doi.org/10.1177/02557614251315461>
- Zhang, L. X., Qiu, Y., Dong, J. Y., & Zou, Y. (2025b). Implementation strategies and challenges of formative assessment: An interview study of primary and secondary school music teachers. *China Music Education*, 3, 5–14. <https://d.wanfangdata.com.cn/periodical/zgyjy202503001>

Author biographies

Le-Xuan Zhang is a research assistant professor at the Department of Cultural and Creative Arts, the Education University of Hong Kong (EdUHK). Dr. Zhang has years of experience working in school teaching and leadership roles and has acted as a K–12 teacher trainer, sharing innovative pedagogical strategies and methods with local and international teachers in China. Dr. Zhang's research interests are in music education, student-centered education, integrated curriculum (e.g., interdisciplinary and transdisciplinary), and educational assessment in authentic learning contexts.

Zi Yan is a senior research fellow of the Hong Kong Research Grants Council and a full professor in the Department of Curriculum and Instruction at the Education University of Hong Kong. His publications and research projects focus on two related areas: educational assessment in school and higher education contexts, with an emphasis on student self-assessment, and Rasch measurement, particularly its application in educational and psychological research.

Xiang Wang is a PhD student of the Department of Sport, Physical Education and Health at the Hong Kong Baptist University and was previously a research assistant in the Department of Curriculum and Instruction at the Education University of Hong Kong. His publications and research interests focus on two areas: meta-analysis in educational and psychological research and sports psychology with an emphasis on promoting athlete mental health.