



## From policy to practice: a longitudinal interview study of experienced music teachers on implementing formative assessment in China

Le-Xuan Zhang, Yue Qiu, Ying Zou & Zi Yan

**To cite this article:** Le-Xuan Zhang, Yue Qiu, Ying Zou & Zi Yan (23 Jan 2026): From policy to practice: a longitudinal interview study of experienced music teachers on implementing formative assessment in China, Music Education Research, DOI: [10.1080/14613808.2026.2615436](https://doi.org/10.1080/14613808.2026.2615436)

**To link to this article:** <https://doi.org/10.1080/14613808.2026.2615436>



© 2026 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 23 Jan 2026.



Submit your article to this journal [↗](#)



View related articles [↗](#)



View Crossmark data [↗](#)

# From policy to practice: a longitudinal interview study of experienced music teachers on implementing formative assessment in China

Le-Xuan Zhang <sup>a</sup>, Yue Qiu <sup>b</sup>, Ying Zou <sup>c</sup> and Zi Yan <sup>a</sup>

<sup>a</sup>The Department of Cultural and Creative Arts, The Education University of Hong Kong, Hong Kong, Hong Kong;

<sup>b</sup>School of Educational Science, Chengdu Normal University, Chengdu, People's Republic of China; <sup>c</sup>Shenzhen Longgang Foreign Language School (Group) Yunhe School, Shenzhen, People's Republic of China

## ABSTRACT

Formative assessment, rooted in constructivist and student-centred theory, is increasingly promoted in Chinese curriculum policies, yet its implementation in Chinese music education remains under-researched. This study addresses that gap through a longitudinal qualitative design, following twelve motivated public school music teachers in a well-resourced region of China. Based on two rounds of interviews conducted over two years, it explores shifts in teachers' perceptions and practices within a culturally embedded educational setting. Using Black and Wiliam's five formative assessment strategies as the analytical lens, the study finds that while teachers increasingly adopt policy-aligned language, their understanding often remains superficial. Strategies are adapted to meet classroom realities, with feedback primarily used to adjust teaching rather than to improve learning. Peer- and self-assessment remain rare, hindered by concerns over classroom management and student readiness. Rather than resistance or failure, the policy-practice gap reflects a cultural reinterpretation of formative assessment. The study highlights tensions between Western-informed assessment models and China's collectivist, exam-oriented system. Teachers' growing interest in practical professional development, such as demonstration lessons, reveals the limitations of abstract, top-down reforms. The study calls for culturally responsive, context-sensitive theory building and intervention-based research to support meaningful pedagogical change in Chinese music education.

## ARTICLE HISTORY

Received 16 January 2025


Accepted 27 December 2025

## KEYWORDS

Formative assessment; China music education; music teacher; curriculum standards; interview

## Introduction

School music education in China operates within a unique teaching context characterised by large classes of 40–55 students, standardised curricular requirements, and hierarchical accountability systems (Zhang and Leung 2023). Into this environment, Western-oriented student-centred education (SCE) reform, from the perspective of constructivism (Fosnot 2013), and its aligned principle of formative assessment, was officially introduced in 2001 and reaffirmed in subsequent policies, including the latest 2022 Arts Curriculum Standards (ACS). However, a fundamental philosophical tension has defined the last two decades of reform. While Western SCE presupposes individual autonomy, dialogic feedback, and dedicated time for peer- and self-assessment, its implementation

**CONTACT** Le-Xuan Zhang  zlxuan@eduhk.hk

© 2026 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

in China has resulted in a collective interpretation that prioritises whole-class engagement and teacher-led performance demonstrations (Zhang, Leung, and Yang 2022).

Nowhere is this policy-practice gap more evident than in formative assessment, which a large-scale survey confirmed remains insufficient in music classrooms (Zhang, Yan, and Wang 2025). This raises an urgent question: *how do teachers actually perceive and apply formative assessment within this reality?* To answer this, our observational longitudinal study uses an interview-based approach to explore the practices of twelve experienced, well-resourced music teachers before and after the 2022 ACS release. The focus on these teachers is intentional: if implementation proves challenging even for the highly motivated and better supported teachers, it might point to profound systemic barriers that policy alone cannot overcome. This study therefore aims to illuminate the persistent disconnect between macro-level policy and micro-level classroom reality, providing vital insights for bridging the divide.

## Formative assessment

Formative assessment is an interactive process whereby teachers, learners, or peers gather, interpret, and use evidence of student learning to guide the next steps in teaching and learning (Black and Wiliam 1998). Whereas summative assessment evaluates learning at the end of a unit through grades or tests, formative assessment with its feedback loop focuses on improving learning as it occurs (Hattie and Timperley 2007). Black and Wiliam (2009) outlined five key teaching strategies for implementing formative assessment: (1) clarify and share learning intentions and success criteria, (2) engineer discussions and tasks to gather evidence of understanding, (3) provide feedback that advances learning, (4) empower students to assess peers, thereby providing instructional resources for each other, and (5) foster students' self-assessment to give them ownership of their learning. These strategies promote dynamic teacher–learner interactions, fostering self-regulation and autonomy (Yan et al. 2023). They particularly create 'moments of contingency' (Black and Wiliam 2009, 10) during which teachers adjust instruction based on real-time evidence, supporting deeper learning and long-term development.

Research has consistently shown that formative assessment enhances student learning. It promotes evidence-based instruction and enhances metacognitive skills, encouraging students to reflect on and take ownership of their progress (Yan et al. 2023). Formative feedback nurtures a growth mindset by emphasising effort and improvement over fixed abilities (Yan, King, and Haw 2021). However, its effective implementation depends on the context, such as classroom culture, teacher expertise, and instructional design (Black and Wiliam 2009). Yan et al. (2021) highlighted four key contextual factors: internal school support, external policy, the school environment, and working conditions. Zhang, Yan, and Wang (2025) found that cultural and contextual differences shaped formative assessment intentions and practices in music teaching. These findings underscore the importance of aligning formative assessment with the educational culture and context.

## Curriculum reform and formative assessment policy in Chinese music education context

Formative assessment in China's school music education has evolved within a highly centralised policy system. While national standards provide clear direction, their implementation often clashes with cultural norms, such as exam-driven learning, teacher authority, and large class sizes, that limit classroom flexibility (Zhang and Leung 2023). Though this top-down approach aims for fairness and standardisation, it poses practical challenges for formative assessment to take root.

The concept of formative assessment first entered China's basic education discourse in 1994, introduced to complement summative traditions by promoting a more holistic view of student progress (Jia 2012). In 1996, Zhou proposed a 'meta-formative' model that positioned formative

assessment as equally important as summative methods. This idea resonated with China's long-standing emphasis on rigorous, diverse evaluation (Liu 2003; Zhou 1996), and soon extended into various subject areas, including music.

In the music domain, formative assessment was formally endorsed in the 2001 Music Curriculum Standards (MCS). While the policy began encouraging student-centred learning, music education still prioritised technical skills and measurable outcomes. The MCS emphasised that evaluation should be clear, practical, and embedded in instruction, advocating a blend of formative and summative approaches, alongside self-, peer-, and teacher-assessment to support student growth (Ministry of Education 2001). Yet, deeply rooted content-focused and teacher-led traditions made it difficult to fully embrace process-based assessment in everyday classrooms (Zhang and Kamarudin 2022).

A decade later, the 2011 revision of the MCS marked a shift from skill acquisition to music literacy development, emphasising comprehensive competencies to be developed and assessed progressively (Ministry of Education 2011). Although the overall stance on formative assessment remained, an important addition clarified its dual function: to affirm students' progress, effort, and achievement, and to identify their learning challenges and improvement strategies – thus supporting their overall development. This mirrors Black and Wiliam's (2009) view of formative assessment as continuous, classroom-based learning support. However, research shows that implementation remains limited. Peer- and self-assessment are still rare, while teacher-directed methods and high-stakes testing continue to dominate (Zhang, Leung, and Yang 2022).

The 2022 upgrade to the Arts Curriculum Standards (ACS) deepened the commitment to formative assessment. It reframed assessment goals around 'core literacy', placing greater emphasis on whole-person development (Ministry of Education 2022). The term 'feedback' appeared more frequently in the standards, and clearer expectations were set for students to evaluate creative work through teacher-, peer-, and self-assessment, offering constructive feedback for future improvement, closely aligning with Black and Wiliam's (2009) principles.

However, despite this steady policy evolution and growing advocacy from educators, empirical studies on classroom-level implementation remain limited. After decades of reform, it still remains unclear how teachers' intention and implementation of applying formative assessment in everyday music teaching practices. Regarding this, the lead author conducted a formative assessment project<sup>1</sup> to fill in the gap. This study reports the second stage of this project.

### **Formative assessment project and the first stage study**

This two-stage project investigated music teachers' perceptions and implementation of formative assessment in Mainland China, with a focus on understanding its function, feasibility, and application in music education. The first stage was a quantitative study grounded in the Theory of Planned Behaviour (Zhang, Yan, and Wang 2025). This stage involved the collection of questionnaires from 671 music teachers across 29 provinces and cities in Mainland China to examine their intentions to use formative assessment, their current practices, and the factors influencing both. Results revealed that while formative assessment practices have not yet been widely implemented, music teachers with positive attitudes and high levels of confidence are more likely to intend to engage in formative assessment. Furthermore, the teachers' intention to use formative assessment mediates the relationship between their attitudes and their actual practices, suggesting that the stronger a teacher's commitment to formative assessment, the greater the likelihood of its use in the classroom.

However, the first stage study is limited in its explanatory power. It reveals relevance but does not explain how teachers translate their intentions into specific classroom practices or why certain factors are influential. Furthermore, it did not account for the top-down influence of curriculum reform and policy within China's hierarchical education system (Zhang and Leung 2023). While

we know that stronger intentions predict greater implementation, the origins of these intentions and the persistent challenges that even motivated teachers encounter remain unexplored.

To address these gaps, a second-stage study is essential. Given that professional development in China often relies on learning from ‘model teachers’ (Zhang, Yan, and Wang 2025), this research focuses on understanding the perceptions of a group of educators who have better resource support and have demonstrated a strong intention and opportunity to implement formative assessment. Their insights are crucial for uncovering the practical realities of translating policy into practice.

## Purpose of study

This paper presents the project’s second stage, a qualitative study designed to gain an in-depth understanding of experienced music teachers’ perceptions of formative assessment. Adopting a longitudinal interview approach (Gubrium 2012), this study tracked these perceptions over a two-year period surrounding the release of the 2022 Arts Curriculum Standards (ACS). Guided by Black and Wiliam’s (1998) formative assessment framework, the inquiry is structured around three central research questions (RQs).

RQ1: What are the perceptions of a group of well-resourced and highly motivated music teachers regarding the purpose and practice of formative assessment in music education?

RQ2: What strategies do these teachers use to implement formative assessment in their classrooms?

RQ3: How did these teachers’ perceptions and practices of formative assessment evolve or remain consistent in the period before and after the release of the 2022 ACS?

## Methodology

This study employs an interview-based approach (Gubrium 2012) within an observational longitudinal design (Caruana et al. 2015). Interviews are a critical method for gathering in-depth insights into participants’ experiences, beliefs, and narratives, allowing researchers to explore complex social phenomena that quantitative methods may overlook. The observational longitudinal design involves tracking the same participants over time without any experimental intervention. This approach is crucial for our research as it enables us to map the evolution of this group of highly-motivated and well-resourced teachers’ perceptions and practices as they naturally occur in response to a policy shift. By combining these methods from a qualitative interpretative perspective, this study captures a deep understanding of this group of teachers’ perceptions and implementation of formative assessment before and after the release of the 2022 ACS in China, thereby supporting an exploration of the transition between policy initiation and classroom practice. The lead author secured ethical clearance from her university’s human research ethics committee. The participants consented to their anonymised data being collected, used, and published in this article.

## Participants

Participants for this qualitative study were selected from the initial survey sample using purposive sampling (Brinkmann 2013). The inclusion criteria required that participants must:

- (1) Be an in-service school music teacher who had participated in the initial survey and indicated a willingness to join the follow-up interview study;
- (2) Show a strong intention in the survey for implementing formative assessment in their teaching;
- (3) Possess a minimum of four years of teaching experience within the public school system; and
- (4) Have previously received formal training in Student-centred Education (SCE) and/or assessment/evaluation approaches.

The sampling strategy also aimed to ensure the final group covered both elementary (Grades 1–6) and middle school (Grades 7–9) levels. High school teachers (Grades 10–12) were excluded from this study, as their students have significantly less instructional time in music due to preparation for the national university entrance examination ('Gaokao').

Following a review of the survey responses against these criteria, 12 in-service music teachers were recruited. The final sample had teaching experience ranging from 4 to 22 years. Critically, all participants were associated with a district-funded educational studio in City Z, which provides them with regular professional development in student-centred education and curriculum reform. Their teaching responsibilities covered Grades 1 through 9 and they participated in both rounds of interviews over the two years. Specific demographic information about the participants can be found in Table 1.

### **Procedure and data collection**

This study employed a longitudinal design with two rounds of semi-structured interviews (Brinkmann 2013). The first round was conducted in March 2022, immediately prior to the official release of the 2022 Arts Curriculum Standards (ACS). The second round was conducted in May 2024, allowing for a two-year period for teachers to engage with the new standards. All 12 participants completed both rounds of interviews. Each interview lasted approximately 40–50 min and was conducted via the online platform WeChat to accommodate the participants. With participant permission, all interviews were audio-recorded and transcribed verbatim for analysis. Following transcription, each participant was invited to review their transcript to ensure the accurate capture of their intended meaning (a process known as member checking). Three major interview questions (supported by 8–10 s-level questions) were asked in each round:

First round:

- (1) How do you perceive formative assessment in music teaching?
- (2) What strategies have you applied to implement formative assessment?
- (3) What challenges or other influencing factors have you encountered when applying formative assessment in teaching?

Second round: 'After the release of ACS ...'

- (1) Have you perceived formative assessment in teaching differently?
- (2) Have you changed or retained your teaching strategies regarding formative assessment?
- (3) What challenges or other influencing factors have you still encountered when applying formative assessment?

**Table 1.** Demographic data ( $N = 12$ ) on participants.

Music teacher (MT)	Teaching grade(s)	*Reporting grade levels	Years of teaching	No. of students	Frequency of teaching at reporting levels (per week)	**No. of classes/week
MT1	1, 2, 7, 8	MS	14	45–50	Once	12
MT2	3, 6, 7	MS	14	50–55	Once	16
MT3	1, 2	ES	8	50	Twice	20
MT4	5, 6, 7, 8, 9	MS	14	50–55	Once	22
MT5	3, 4, 5	ES	12	50–55	Twice	15
MT6	5, 8	MS	22	40–55	Once	18
MT7	1, 2	ES	9	45	Twice	19
MT8	1	ES	9	50	Twice	19
MT9	1, 3	ES	5	47	Twice	20
MT10	1, 6	MS	12	50–55	Once	17
MT11	1, 3	ES	4	50	Twice	16
MT12	5, 7, 8, 9	MS	4	50–70	Twice	14

\*Reporting levels: MS – middle school; ES – elementary school.

\*\*No. of classes/week: includes regular classroom teaching (40 min/each) and after-school music club lessons (1–1.5 h/each).

## Data analysis

Our data analysis was systematically structured around the two interview cycles to rigorously track the evolution of teacher perceptions over the two-year period. The process was designed to first establish a baseline and then identify and interpret changes over time.

1. Data Source and Analytical Phasing by Interview Cycle – Our analysis was phased to correspond with the two interview cycles, allowing us to first establish a baseline and then track changes over time.
  - *Phase 1 (Initial Interviews analysis, July-August, 2022)* – This analysis established a baseline understanding of teachers' views and practices before the 2022 ACS release, directly addressing RQ1 (perceptions) and RQ2 (strategies).
  - *Phase 2 (Follow-up Interviews analysis, July-September, 2024)* – This analysis focused on the period after the ACS release. It was designed to investigate the evolution of teacher perceptions and practices, thereby answering RQ3 (evolution/consistency).
2. Thematic Analysis Approach – We employed a two-cycle thematic analysis (Saldaña 2021) to systematically develop themes from the data.
  - *First Cycle Coding:* This initial stage used a hybrid strategy to organise the raw data for analysis. Data for RQ1 and RQ3 was coded inductively with In Vivo and Pattern Coding. For RQ2, data was coded deductively against Black and Wiliam's (2009) five formative assessment strategies using descriptive coding.
  - *Second Cycle Coding:* This stage focused on synthesis of the organised data into conceptual categories. Focused coding was applied in the second cycle of coding to identify significant codes that formed our final themes and subthemes (see Table 2).

## Trustworthiness

To establish the trustworthiness of this qualitative study, we implemented strategies addressing the four criteria: credibility, dependability, conformability, and transferability (Haq et al. 2023).

First, to ensure dependability, a structured coding manual was developed to create a clear audit trail. The data were divided into three sections, *pre-ACS perceptions*, *post-ACS perception*, and

**Table 2.** Examples from two cycles of data analysis.

Raw Data	First Cycle Coding		Second Cycle Focus
	In Vivo	Pattern	
My student needs to participate in high-stake test ... I don't have time. (MT6)	Lack of time for formative assessment due to student's test preparation	Anxious about time	Challenges and concerns in teaching
The most difficult thing is to find time for students to assess. (MT12)	Finding time for assessment		
I need to be concerned about time because when students get too excited, they cannot finish the tasks I assign. (MT8)	Concern about time		
I don't have energy to 'research' ACS or make my teaching [assessment] better. (MT5)	Lack of energy due to heavy workload	Lacking energy	
Priority concern is the teaching loads ... teaching like a robot and no energy for others. (MT7)	Teaching load leaves no energy for others		
Peer assessment takes time ... spending time telling them criteria and gathering their feedback takes a lot of energy. (MT8)	Peer assessment drains energy		
I attended some training, but I felt tired afterward. (MT3)	Participating in training is tiring	Burnout	
The teaching load is heavy, so I feel tired. (MT7)	Teaching load is tiring		
Too much non-teaching work causes me annoyed. (MT7)	Non-teaching work annoys me		

*implemented strategies*, with the first two coded inductively and the implementation section coded deductively using Black and Wiliam's (2009) framework.

Second, we employed multiple strategies to enhance credibility. The initial coding was conducted independently by the second author. These codes and interpretations were then reviewed and discussed during weekly online discussions with the first author. Both authors are university researchers with extensive experience in Chinese arts education, the collaborative process served as a form of peer debriefing. The third author, with more than 10 years music teaching experiences in Chinese public school, performed member checking with the interviewees to validate that our interpretations accurately reflected their experiences.

Third, to ensure conformability and reduce researcher bias, we applied investigator triangulation. This was achieved through the collaborative analysis of researchers with diverse perspectives and, critically, through an expert review by the fourth author, a university professor and research expert in formative assessment. This expert review also contributed significantly to the study's credibility and dependability.

Finally, by grounding our analysis in the participants' both round perceptions, we are able to provide a thick description in our findings, allowing readers to assess the potential relevance of the findings to their own contexts.

## Results

### *Teachers' perceptions of formative assessment*

Generally, participating teachers generally perceived formative assessment as a valuable practice, articulating a multifaceted understanding of its functions while also acknowledging significant limitations in their own knowledge. Their perceptions centred on three key areas: the varied interpretations of the practice's function, the self-reported challenges stemming from inadequate training, and the resulting shift in their professional roles as educators.

First, regarding the interpretation of formative assessment, teachers most frequently identified its function as a *diagnostic tool for tracking student learning*. They considered it crucial for understanding students' current learning stage, citing the need for 'evaluations at multiple mid-points to understand student progress' (MT1) and for 'understanding students' in-process musical expression and performance' (MT6), and for improving instruction. As MT10 explained, this allows for timely intervention: 'Formative assessment can help me receive prompt feedback ... I can fill in the water if I find the full bucket has lost  $\frac{1}{4}$  of its water'. In addition, some teachers emphasised its function as a *motivational teaching strategy*. For instance, MT9 noted that using simple rewards like stamps makes students 'willing to answer questions', which in turn 'stimulates students' classroom participation and proactivity'. Moreover, a more nuanced interpretation was offered by MT11, who described it as a *co-constructed, interactive process* 'of joint exploration between the teacher and students', rather than something 'completed by the teacher alone'.

However, beyond these varied interpretations, 10 teachers reached a consensus: they possessed only a superficial understanding of the concept and lacked the skills for its practical implementation (MT2, MT3, MT4, MT5, MT7, MT8, MT9, MT10, MT11, MT12). For example, MT2 admitted, 'I heard about it, but I did not have an in-depth understanding of it'. MT4 echoed this sentiment, stating, 'I cannot clearly articulate what formative assessment is ... I feel stuck at a superficial understanding of the terminology without grasping the deeper concepts'. Similarly, MT3 highlighted the gap between awareness and application, noting, 'I am aware of it, but I don't know the specific implementation approach'. Reinforcing these points, and referencing similar comments from MT8, MT9 highlighted the disconnect between official requirements and classroom reality: 'I know formative assessment is a major component in our Teaching Reference Book<sup>2</sup> ... and we are required to use it. But because of a lack of deep understanding, I personally still do not use it that much'.

Finally, seven teachers reported that engaging with formative assessment prompted a significant shift in their professional role (MT1, MT2, MT4, MT7, MT8, MT9, MT12). They described moving from a traditional, teacher-centred model to becoming guides, facilitators, and motivators. As MT1 explained, this new role is akin to being a navigational tool: ‘When using formative assessment, teachers are like compasses or GPS devices, guiding students towards learning objectives based on their progress ... This enables teachers to adjust their subsequent teaching strategies and plans’. Both MT4 and MT9 contrasted this new approach with the traditional model, where a teacher ‘is just ‘outputting’, not knowing how the students are learning’. In this new model, teachers are required to ‘change their teaching approach’ based on real-time student feedback. This results in more responsive and flexible teaching practices and ensures better alignment with curriculum standards that mandate a ‘focus on individualised student needs’.

### **Formative assessment strategies implemented**

Although teachers reported a superficial understanding of formative assessment and limited practice, they were still able to share how they applied their own interpretations of it. Given that the first stage survey revealed a lack of formative assessment practices among Chinese music teachers (Zhang, Yan, and Wang 2025), this analysis does not aim to inductively generate new findings from their responses. Instead, following the established framework of Black and Wiliam (2009), this section deductively presents how five formative assessment strategies have been implemented with teachers’ interpretation and variation in this collective SCE-based Chinese educational context (Zhang and Leung 2023), allowing for comparison with Western assessment contexts.

Overall, the results showed that Teachers primarily applied Strategies 1 and 3, used classroom discussion (S2) less, and rarely employed peer (S4) or self-assessment (S5). Crucially, their practices were not purely for formative assessment; they consistently blended it with techniques for classroom management, motivation, and contextualisation. Thus, this analysis uses Black and Wiliam’s framework (2009) to reveal how teachers actually perceive and use assessment, showing these functions to be inseparable in their practice.

#### **Strategy 1 (S1) – clarifying learning goals and criteria**

All teachers commonly mentioned to provide clear objectives, guide students’ efforts, and ensure purposeful engagement in their teaching practices. Regarding formative assessment, as this strategy is primarily teacher-led, they focused on explaining how they set teaching goals to help students understand the subsequent learning content. For example:

If I teach a 3/4-time signature, I will sing the ‘Hello’ song with students in 3/4 time, ask them to respond to me in the same 3/4 time and share what I intend for them to do. This helps them better understand and respond with the appropriate rhythm. (MT5)

I set clear teaching goals before each lesson and aim to spend 10 min extending students’ knowledge in every class. For example, once, I taught my students a famous Chinese song, ‘Tian Lu’, describing the impressive feat of Chinese people building a railway line connecting Qinghai and Tibet. I prepared background information for my students to help them understand the song before learning it ... (also) help them to get an impression of learning and song with its behind resilience and Chinese culture meaning. (MT6)

I tell them what they must do and praise them when they complete it well. They perform better when they understand the learning goal and feel proud after I commend them for achieving it, such as maintaining good posture, showing expressive facial expressions, or performing movements effectively. (MT8)

These examples demonstrate the varied ways in which teachers approach the clarification of learning goals. Their practices include not only defining the technical task (MT5), but also providing cultural context (MT6) and using praise to highlight performance criteria (MT8). These examples show that even the foundational strategy of clarifying goals is interpreted through a lens of student management. The focus is on ensuring compliance and engagement through clear instructions

(MT5), extrinsic motivation via praise (MT8), and contextual framing (MT6), rather than on co-constructing success criteria to foster student autonomy. Thus, from the outset, the formative goal of making learning transparent is intertwined with the pragmatic need to manage and motivate a large class.

### **Strategy 2 – engineering effective classroom discussions to gather learning evidence**

While formative assessment literature emphasises rich classroom discussion for uncovering student thinking, this strategy was absent from the teachers' descriptions. When asked about this absence, teachers cited two barriers: a lack of confidence in students' discussion ability (MT4, MT5) and concerns about classroom management (MT6).

I know my assessment should be based on them doing these tasks, then discussing and sharing with each other. However, what can I do if they don't speak or share? Yes, what can I do? I don't know what to do ... I don't think lower-grade students can truly do this. (MT4)

This (classroom discussion) happens rarely – it's very uncommon in Grade 5. But earlier this semester, I did try it once ... I asked the students to pair up ... But I guess they didn't really know how to do it properly. (MT5)

Very occasionally I will organise a class discussion. When a class feels dull or unengaged, I have students write their thoughts in small groups and have a brief discussion. This can definitely get a bit noisy, so I need to set clear expectations: they have three minutes to write seriously, and order must be maintained. To encourage good behavior, the most orderly group gets to present first. (MT6)

These examples show that teachers avoided classroom discussions due to a lack of confidence in both their students' abilities and their own facilitation skills (MT4, MT5). Crucially, even when discussions were attempted, their formative purpose was secondary to classroom management. Teacher MT6, for instance, used a group activity not to gather learning evidence but to correct low engagement, rewarding orderly behaviour rather than intellectual contribution. Thus, the pragmatic goal of managing the classroom supplanted the formative goal of understanding student thinking.

### **Strategy 3 (S3) – providing feedback to move learners forward**

Feedback was mentioned frequently among teachers. However, while much research focuses on the feedback, given *from the teacher to learners*, to support learning (Frank, Simper, and Kaupp 2017), this study revealed that eight teachers interpreted student engagement as feedback, *from the learners to the teacher*, to improve teaching (MT1, MT2, MT4, MT5, MT7, MT9, MT10, MT12).

I observed and gathered students' feedback on my teaching. For example, I let my students use rhythm to create movements. The students provided active feedback and had positive attitudes, so I knew they liked to learn it. But ... when I asked them to learn from the songs in the textbook, the students showed less feedback to me. (MT7)

I take students' feedback to evaluate my teaching content and style. If students are too quiet and do not react actively and positively, I am not able to evaluate their learning progress. So, I have to chat with them individually or call on them by their student number to have them sing or answer questions as feedback on their learning. (MT5)

Although two teachers mentioned providing feedback to students, their examples often focused on whole-class praise to affirm students' creativity rather than asking for another round of improvement to enhance individual students' learning progress (MT2, MT7).

I asked them to listen to Beethoven's 'Tempest' sonata while drawing what they imagined. I checked students' individual drawings and was impressed by their imagination. So, I told them I was impressed ... , because their drawings of wolves, forests, and small animals could provide a richer picture than the piano piece originally presented. (MT2)

These findings reveal a contextual understanding of the feedback. Instead of providing diagnostic advice to students, these Chinese music teachers primarily sought feedback from students to evaluate their own teaching. They interpreted observable student engagement, such as active participation and positive attitudes, as the primary indicator of successful learning. Consequently, feedback functioned less as a tool to advance individual student learning, and more as a real-time gauge for teachers to manage classroom dynamics and adjust their instructional style.

#### ***Strategy 4 – activating students’ peer assessment***

While teachers acknowledged the potential benefits of peer assessment for building confidence and critical thinking (MT8, MT11), they expressed significant hesitation in using it. The primary concerns were that students lacked the skills to assess peers constructively and that the activity could devolve into mockery or classroom chaos (MT6, MT8, MT11). The account from MT6 is particularly revealing, as it shows how these concerns lead to peer assessment being recast as a highly controlled, behaviour-contingent activity rather than a standard learning practice.

If students are active and positive, I may ask them to evaluate each other’s work by separating them into groups and allowing brief evaluations. I set clear guidelines to keep the class manageable: a timer to limit speaking time and a competition mechanism where only well-behaved groups can present their assessments. Having taught music for over two decades, I’ve learned that making disruptive students feel valued and cared for often encourages better behaviour and greater engagement. (MT6)

This evidence reveals that peer assessment was deployed not as a learning tool, but as a conditional reward for good behaviour. The teacher’s method was dominated by strict procedural controls designed to maintain order rather than to ensure high-quality feedback. The primary goal was therefore not to generate formative insights, but to manage the classroom by channelling student activity into a highly structured and competitive format.

#### ***Strategy 5 – fostering students’ self-assessment***

Like peer assessment, self-assessment was rarely used. Several teachers mentioned that they did not use it at all (MT3, MT7, MT11, MT12). The primary reasons given were a lack of student capability and practical classroom constraints. MT8 exemplified the belief that students were not capable of deep self-reflection, noting that the strategy is ‘challenging to implement with younger students’. For example,

I only used it to ask students to briefly evaluate their singing, specifically whether they followed the correct rhythm or melody patterns. I think music learning should focus on technique, and issues such as incorrect rhythm or off-pitch singing are the easiest for students to identify and improve on. (MT8)

Thus, this evidence, aligning with findings on classroom discussion (Strategy 2), shows that teachers lack confidence in students’ assessment abilities (including discussion and self-evaluation). This lack of confidence, coupled with concerns about assessment accuracy and the loss of teaching time, makes self-assessment difficult to implement. Because of these practical challenges, the strategy was either not used or was simplified to a basic check on rhythm and melody. Consequently, it was treated as an optional activity rather than an essential part of the lesson.

### ***Changes and continuities***

The longitudinal analysis, conducted over two years spanning the release of the 2022 ACS, reveals a complex narrative of superficial adaptation, deeply entrenched barriers, and a crucial evolution in teachers’ professional awareness.

#### ***A Superficial shift: adopting new language without deepening understanding***

The most immediate change observed between the two interview rounds was the teachers’ adoption of policy-aligned language. In the second interview, teachers more frequently used student-centred

terminology, reflecting the emphasis of the new ACS. For example, eight teachers (MT1, MT2, MT3, MT4, MT5, MT6, MT11, and MT12) reported a renewed focus on the student, with MT2 explaining that evaluation should move beyond mere skills to encourage ‘confident self-expression’.

However, this linguistic shift masked a persistent and profound conceptual gap. A critical finding that remained consistent across both interviews was the teachers’ fragmented and shallow understanding of formative assessment. In the second round, this was articulated even more sharply. Seven teachers (MT2, MT3, MT4, MT5, MT6, MT7, and MT10) admitted their knowledge was largely theoretical, a gap powerfully captured by MT3: ‘I still don’t have a clear theoretical knowledge of it. Therefore, if I practised it, I might not even know that I had’. This sentiment was echoed by MT4, who dismissed the policy changes as merely ‘different wording’ for existing practices. This evidence suggests that while the macro-level policy prompted a change in what teachers said, it might fail to fundamentally change how they understood assessment at a micro-level.

### ***Persistent barriers: the enduring conflict between pedagogical aspiration and managerial reality***

The most striking continuity across the two years was the unresolved conflict between teachers’ pedagogical goals and the managerial realities of their classrooms. This theme was prominent in the first round of data (coded under challenges like ‘Large class sizes’ and ‘Excessive workload’) and remained a primary barrier in the second round.

While teachers expressed a genuine desire to implement student-led activities, this aspiration was consistently overridden by a fear of losing classroom control, a sentiment captured by the code ‘With student-centredness, classroom management is difficult’. Seven teachers (MT1, MT3, MT4, MT5, MT7, MT8, and MT10) associated student-led learning with ‘loud and disruptive’ chaos. This fear had not diminished by the second interview; if anything, it had become more entrenched under the pressure to meet new curriculum targets. MT12 articulated this persistent anxiety perfectly: ‘I would be afraid to hand the classroom over to them’.

This internal conflict was exacerbated by external barriers that were cited with equal frequency in both interview rounds. Large class sizes, often exceeding 50 students, made individualised strategies feel impractical (MT4, MT7). Heavy workloads drained the time and energy required for experimentation, with MT7 describing feeling ‘like a robot’. When overwhelmed, half the teachers (MT3, MT5, MT6, MT7, MT8, and MT12) admitted they defaulted to more manageable, traditional assessments. The longitudinal data shows that these powerful contextual factors created a reality where managerial concerns consistently overrode pedagogical aspirations, rendering the new policy largely ineffective in changing core classroom dynamics.

### ***An emerging awareness: the growing demand for practice-oriented support***

While core practices and deep understanding saw little change, a notable longitudinal evolution occurred in the teachers’ articulation of their professional needs. In the first round of interviews, their desire for training was present but often general. By the second round, however, having grappled with the policy’s abstract demands for two years, their requests became overwhelmingly specific and urgent.

This shift is a direct result of their practical struggles. The demand for observable models of practice became the dominant theme in the second interviews, as coded under ‘Expectations for Training Content’. Teachers were no longer satisfied with theory. MT11 stated with newfound clarity, ‘I most urgently want to use a really good classroom case study to guide us’. This call for ‘demonstration lessons’ and ‘practical application’ was echoed by at least seven other teachers (MT4, MT5, MT6, MT7, MT10, MT11, and MT12). This finding demonstrates a key longitudinal shift: while the policy failed to change practice directly, it succeeded in making teachers acutely aware of the systemic gap between abstract mandates and the concrete, practice-oriented support they truly needed to bridge the policy-practice divide.

## Discussion

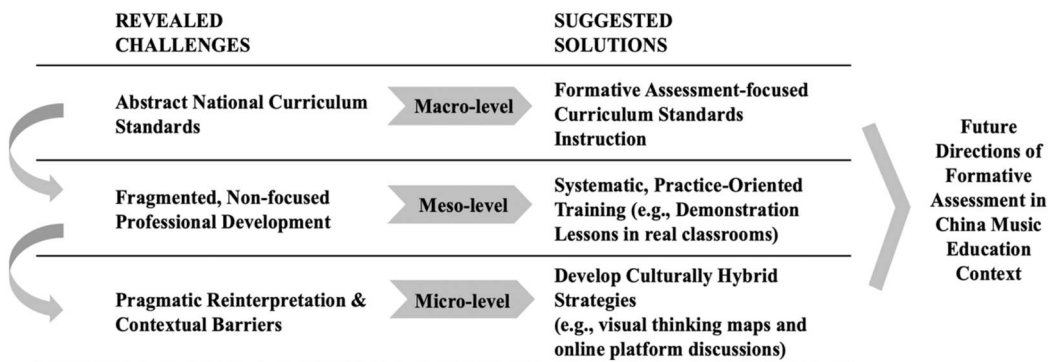
This study explores how 12 music teachers in one of China's most developed cities, each with high-motivated, well-resourced and SCE-training experienced background, perceive formative assessment in public school teaching. The findings suggest five key points: (1) teachers value formative assessment but lack the confidence and theoretical knowledge to apply it effectively; (2) formative assessment strategies are primarily used to support classroom management and student engagement; (3) peer- and self-assessment, emphasised in national policy documents such as the Arts Curriculum Standards (ACS), remain challenging to implement; (4) the release of the ACS in 2022 has not led to significant changes in these teachers' formative assessment practices; and (5) teachers express a strong need for practice-oriented professional support (e.g. demonstration lessons).

These findings reveal phenomena consistent with previous research on SCE in the Chinese music education context (Zhang and Leung 2023), such as the challenges of fully adopting a Western-oriented educational approach and the need for professional development. This study contributes new, assessment-based insights by examining the understanding and practices of teachers who were presumed to have a high likelihood of implementing formative assessment in their lessons. To provide readers with a better understanding of the Chinese music education context and to maintain focus on the gap between policy and practice, the discussion section adopts the same systematic framework used in the research of Zhang and Leung (2023), outlining the macro-, meso-, and micro-levels of China's school music education system. A summary of the key challenges and future suggestions for implementing formative assessment is first presented in Figure 1. More details were be discussed as follow.

### *The macro-level policy: will curriculum reform lead to successful implementation?*

Despite decades of national curriculum standards promoting formative assessment (Ministry of Education 2001, 2011, 2022), our findings reveal that even experienced teachers with strong intentions and better resource support struggle with its implementation, echoing previous research (Zhang, Yan, and Wang 2025). This suggests an impression that the top-down curriculum reform, on its own, is highly likely insufficient to drive meaningful pedagogical change in Chinese music education.

Two key factors explain this persistent policy-practice gap. First, translating any educational concept from theory into practice is a slow process requiring deep contextualisation (Fullan 2012). The Chinese music curriculum standards, while conceptually modern, remain abstract guides. They articulate the 'what' (e.g. use peer – and self-assessment) but offer little of the practical 'how', leaving teachers without actionable strategies to navigate their complex classroom realities.



**Figure 1.** A systemic framework for implementing formative assessment in China music education.

Second, and more critically, this study highlights a systemic shortage at the meso-level. As introduced by Zhang and Leung (2023), curriculum standards are a macro-level policy in China music education system. For such policies to be actualised at the micro-level of classroom teaching, they must be translated and supported by a robust meso-level professional development. This study reaffirms that this meso-level link still remains underdeveloped in Chinese music education. Without strengthening this essential bridge, the answer to whether policy reform leads to successful implementation will remain negative.

### ***The meso-level professional development: does regular training help teaching practice?***

Although teachers in this study receive regular professional development, our findings show it is ineffective for implementing formative assessment. This ineffectiveness stems from two primary issues.

First, teachers require systematic training rather than the fragmented, workshop-based model that currently prevails. Complex pedagogical shifts, such as the move to formative assessment, demand deep and sustained engagement, not a superficial overview (Fullan 2012). Without this, teachers risk developing only a shallow understanding of the new concepts, a danger amplified by China's long tradition of teacher-centred pedagogy. The difficulties reported by our participants provide strong evidence for this gap.

Second, teachers show a strong preference for a specific professional development format: the demonstration lesson (Zhang and Leung 2025). This practice-oriented model is effective because it bridges the theory-practice gap, showing how abstract concepts are contextualised in a realistic classroom. This preference also helps explain why participants intuitively blend formative assessment with other teaching strategies, rather than treating it as an isolated activity as conceptualised by some theorists.

Therefore, we argue that the issue is not an absence of training, but a fundamental mismatch in its form and content. We recommend a shift towards systematic, practice-oriented trainings centred on demonstration lessons, happening in a real classroom setting. In this regard, the FAITH website<sup>3</sup>, developed with both theoretical and video-based practical resources, could serve as a valuable model for strengthening teachers' professional development in formative assessment. This approach would not only better support teachers but also strengthen the crucial meso-level of the Chinese music education system (Zhang and Leung 2023).

### ***The micro-level teaching practice: how should we interpret formative assessment in China music education context?***

This study's findings indicate a reinterpretation of Black and Wiliam's (2009) Western-oriented framework for formative assessment, where its theoretical principles are consistently adapted to serve the immediate, practical needs of the Chinese music classroom. This is evident in how teachers repurposed core strategies for pragmatic ends: 'engineering effective classroom discussions' (S2) was reduced to recitation-style questioning to manage attention (MT5), while peer assessment (S4) became a conditional reward for good behaviour rather than a tool for learning (MT6).

These adaptations are rooted in an enduring teacher-centred culture. Given feedback, as one of the key strategies in formative assessment, can improve both learning and teaching (Hattie and Timperley 2007), this reinterpretation offers a key insight: *in this collectivist context, feedback is unconsciously prioritised as a tool to refine teaching*. This focus on the pedagogical function of feedback echoes previous research (e.g. Gaertner 2014), which argues that the role of feedback in improving teaching, not just learning, warrants greater exploration, especially within such collectivist educational contexts.

While these practices can be partly explained by structural factors like large class sizes (Zhang, Leung, and Yang 2022), a deeper analysis reveals a tension between *cultural diversity* and *cultural*

*hybridity*. Cultural diversity emphasises preserving unique educational traditions and valuing differences rooted in specific social and cultural contexts (Banks 2015). Cultural hybridity involves blending global and local elements to create shared practices and foster mutual integration (Bhabha 2012). An overemphasis on preserving local traditions risks rejecting beneficial global innovations. The path forward, therefore, lies not in abandoning formative assessment, but in embracing cultural hybridity to develop culturally responsive strategies. Chinese researchers have recommended that students use writing/drawing tools to express their thoughts (Zhang and Leung 2023). Digital applications to interact with each other online rather than engaging in oral peer communication has also been recommended with class sizes exceeding 50 (Chen, Jianli, and Kexin 2021). Such hybrid approaches can integrate formative principles into the classroom without causing disruption, creating context-sensitive solutions that are meaningful within China's unique educational reality.

## Conclusion and future direction

This study set out to investigate the gap between formative assessment policy and practice from the perspective of constructivist, student-centred education within a collective-based Chinese teaching context. Our findings conclude that this gap is not an empty space of failed implementation, but rather a site of dynamic cultural reinterpretation.

This leads to an important theoretical and reflective question: Should Chinese educational policies engage more deeply with Chinese traditions and classroom realities, rather than aligning primarily with Western constructivist standards? Considering the uniquely emerging collective SCE context in Chinese music education (Zhang and Leung 2023; Zhang, Leung, and Yang 2022; Zhang, Yan, and Wang 2025), along with this study's key finding that feedback is often unconsciously prioritised as a tool to refine teaching rather than to enhance student learning, it may be time to investigate the contextualised implementation standards and theoretical foundations for music education in China. In line with Bremner's (2020) suggestion of multiple interpretations of SCE, a valuable direction for future research is to provide context-sensitive theoretical contributions that better reflect local pedagogical traditions and values.

Despite its contributions, this study has limitations. Although we applied a collaborative coding strategy involving four authors in different roles throughout the process, we acknowledge that incorporating an inter-rater reliability check could have further enhance the credibility of the analysis. Besides, the small, purposively selected sample of all-female teachers with a strong interest in formative assessment, while a valuable qualitative complement to previous survey work (Zhang, Yan, and Wang 2025), cannot represent all Chinese music teachers. Future research should therefore include more diverse samples, encompassing male teachers, educators with varying experience levels, and those from a wider range of schools.

Looking forward, a critical next step is to design and conduct intervention studies to enhance both teacher and student assessment literacy. Since assessment literacy is foundational to success, research must investigate how it can be cultivated within China's specific cultural and pedagogical context. The ultimate goal is to adapt diverse formative assessment strategies, to align with the unique needs of music education in China, generating insights that are both locally grounded and globally relevant

## Notes

1. *An Investigation of Music Teachers' Intention and Implementation* (RG 19/2023-2024), an internal research grant from the Committee on Research and Development by The Education University of Hong Kong.
2. Teaching Reference Book (Jiao Can): A teacher guide for providing teachers with a detailed, lesson-by-lesson guide on how to effectively teach the content of the textbook.
3. FAITH refers to the *Formative Assessment Innovative Teaching Hub*. It is a website developed to enhance school teachers' professional development in formative assessment. <https://faith.eduhk.hk/>.

## Disclosure statement

No potential conflict of interest was reported by the author(s).

## Funding

This research study was supported by the Start-up Research Grant (number RG 19/2023-2024R) from The Education University of Hong Kong.

## Notes on contributors

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

*Yue Qiu* is a lecturer in the Early Childhood Education Department at the College of Education and Psychology, Chengdu Normal University. She held her doctoral degree from The Education University of Hong Kong and her master's degree from Beijing Normal University. Dr. Qiu has trained thousands of kindergarten teachers. In addition to her role as a college educator, she also teaches children sand painting and drama in her spare time. Her research interests include creative arts education and teacher education.

*Ying Zou* is a music teacher at Shenzhen Longgang Foreign Language School (Group) Yunhe School, overseeing teacher development and leading her music master teacher studio. A graduate of Wuhan Conservatory of Music, she has presented twice at the ISME World Conference. Her research focuses on applying innovative educational theories in schools.

*Zi Yan* is a professor and the head of the Department of Curriculum and Instruction at the Education University of Hong Kong. Prof. Yan has many years of experience in knowledge transfer, teacher training, and school support. His publications and research projects focus on two related areas: (i) educational assessment in the school and higher education contexts, with an emphasis on student self-assessment; and (ii) Rasch measurement, particularly its application in educational and psychological research.

## ORCID

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

*Yue Qiu*  <http://orcid.org/0009-0009-2855-9352>

*Ying Zou*  <http://orcid.org/0009-0005-4078-7463>

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

## References

- Banks, James A. 2015. *Cultural Diversity and Education: Foundations, Curriculum, and Teaching*. United States. New York: Routledge.
- Bhabha, Homi K. 2012. *The Location of Culture*. London, UK: Routledge. <https://doi.org/10.4324/9780203820551>.
- Black, Paul, and Dylan Wiliam. 1998. "Assessment and Classroom Learning." *Assessment in Education: Principles, Policy & Practice* 5 (1): 7–74. <https://doi.org/10.1080/0969595980050102>.
- Black, Paul, and Dylan Wiliam. 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>.
- Bremner, N. 2020. "The Multiple Meanings of 'Student-Centered' or 'Learner-Centered Education', and the Case for a More Flexible Approach to Defining It." *Comparative Education* 57 (2): 159–186. <https://doi.org/10.1080/03050068.2020.1805863>.
- Brinkmann, Svend. 2013. *Qualitative Interviewing*. New York: Oxford University Press.
- Caruana, Edward Joseph, Marius Roman, Jules Hernández-Sánchez, and Piergiorgio Solli. 2015. "Longitudinal Studies." *Journal of Thoracic Disease* 7 (11): E537. <https://doi.org/10.3978/j.issn.2072-1439.2015.10.63>.
- Chen, Zexuan, Jiao Jianli, and Hu Kexin. 2021. "Formative Assessment as an Online Instruction Intervention: Student Engagement, Outcomes, and Perceptions." *International Journal of Distance Education Technologies (IJDET)* 19 (1): 50–65. <https://doi.org/10.4018/IJDET.20210101.oal>.
- Fosnot, C. T.. 2013. *Constructivism: Theory, Perspectives, and Practice*. New York: Teachers College Press.

- Frank, Brian, Natalie Simper, and James Kaupp. 2017. "Formative Feedback and Scaffolding for Developing Complex Problem Solving and Modelling Outcomes." *European Journal of Engineering Education* 43 (4): 552–568. <https://doi.org/10.1080/03043797.2017.1299692>.
- Fullan, Michael. 2012. *Change Forces: Probing the Depths of Educational Reform*. London: Routledge.
- Gaertner, Heidi. 2014. "Effects of Student Feedback as a Method of Self-evaluating the Quality of Teaching." *Studies in Educational Evaluation* 42:91–99. <https://doi.org/10.1016/j.stueduc.2014.04.003>.
- Gubrium, Jaber F. 2012. *The Sage Handbook of Interview Research: The Complexity of Thbke Craft*. 2nd ed. Thousand Oaks, CA: Sage.
- Haq, Zia Ul, Rizwana Rasheed, Aamir Rashid, and Salma Akhter. 2023. "Criteria for Assessing and Ensuring the Trustworthiness in Qualitative Research." *International Journal of Business Reflections* 4 (2): 150–173. <https://doi.org/10.56249/ijbr.03.01.44>.
- Hattie, John, and Helen Timperley. 2007. "The Power of Feedback." *Review of Educational Research* 77 (1): 81–112. <https://doi.org/10.3102/003465430298487>.
- Jia, Junhong. 2012. "Xingcheng xing pingjia de yanjiu xianzhuang yu fazhan qushi - yi Zhongguo xueshu qikan wangluo chuban zongku wei yangben lai yuan" [On the Current Situation and Development Trend of Formative Assessment - Taking the CNKI of China as Sources]. *Journal of Sichuan University of Science & Engineering (Social Sciences Edition)* 27 (1): 106–109. <http://xb.suse.edu.cn/upload/20120210/201202100629162055.pdf>.
- Liu, Rao. 2003. "'Zhongguo jiaoyu pingjia fazhan lishi shuping'" [Review on the Development History of Chinese Educational Appraisal]. *Journal of Beijing University of Technology (Social Sciences Edition)* 3 (3): 88–92. <https://tra-oversea-cn2ki-net.ezproxy.eduhk.hk/knavi/JournalDetail?pcode=CJFD&pykm=BGYS>.
- Ministry of Education, PRC. 2001. *Music Curriculum Standards of Compulsory Schooling (Trial)*. Beijing, China: Beijing Normal University Press.
- Ministry of Education, PRC. 2011. *Music Curriculum Standards of Compulsory Schooling*. Beijing, China: Normal University Press.
- Ministry of Education, PRC. 2022. *Arts Curriculum Standards for Compulsory Education*. 2022 ed. Beijing, China: Beijing Normal University Press.
- Saldaña, Johnny. 2021. *The Coding Manual for Qualitative Researchers*. 4th ed. London, UK: Sage Publications.
- Yan, Zi, Ziqi Li, Ernesto Panadero, Min Yang, Lan Yang, and Hongling Lao. 2021. "A Systematic Review on Factors Influencing Teachers' Intentions and Implementations Regarding Formative Assessment." *Assessment in Education: Principles.* *Policy andamp; Practice* 28 (3): 228–260. <http://doi.org/10.1080/0969594X.2021.1884042>.
- Yan, Zi, Ernesto Panadero, Xiang Wang, and Ying Zhan. 2023. "A Systematic Review on Students' Perceptions of Self-Assessment: Usefulness and Factors Influencing Implementation." *Educational Psychology Review* 35. <https://doi.org/10.1007/s10648-023-09799-1>.
- Yan, Zi, Ronnel B. King, and Joseph Y. Haw. 2021. "Formative Assessment, Growth Mindset, and Achievement: Examining Their Relations in the East and the West." *Assessment in Education: Principles.* *Policy andamp; Practice* 28 (5-6): 676–702. <http://doi.org/10.1080/0969594X.2021.1988510>.
- Zhang, Le-Xuan, and Bo-Wah Leung. 2023. "Context Matters: Adaptation of Student-Centred Education in China School Music Classrooms." *Music Education Research* 25 (4): 418–434. <http://doi.org/10.1080/14613808.2023.2230587>.
- Zhang, Le-Xuan, and Bo-Wah Leung. 2025. "Defining Music Demonstration Lessons: A Unique Performance-Based Lesson Type Improving Teachers' Instructional Skills in Chinese Mainland Education." *British Journal of Music Education* 42 (1): 65–79. <https://doi.org/10.1017/S0265051724000214>.
- Zhang, Le-Xuan, Zi Yan, and Xiang Wang. 2025. "Understanding Music Teachers' Formative Assessment Intention and Implementation: A Chinese Mainland Context." *International Journal of Music Education*. <https://doi.org/10.1177/02557614251315461>.
- Zhang, L.-X., B.-W. Leung, and Y. Yang. 2022. "From Theory to Practice: Student-Centered Pedagogical Implementation in Primary Music Demonstration Lessons in Guangdong." *International Journal of Music Education* 41 (2): 271–287. <https://doi.org/10.1177/02557614221107170>.
- Zhang, Yunchao, and Diyana Kamarudin. 2022. "A Review of Music Education Assessments in China." *International Research Journal of Education and Sciences* 6 (1): 236–242. <https://www.masree.info/wp-content/uploads/2022/12/ICREE-22-87-Zhang-Yunchao.pdf>.
- Zhou, Z. G. 1996. "Dui Zaipinggu de Tantau: Zhide Zhongshi he Yanjiu de Wenti" [Further Discussion on Assessment: The Question Required to be Valued and Researched]. International Education Evaluation and Quality Assurance Conference.