Action, Criticism, and Theory for Music Education
August 2026. Vol 25 (4): 115–38. doi:10.22176/act25.4.115

Decolonizing Aural Skills Assessment in Higher Music Education: Epistemic Authority, Evidence, and Judgment

Eugene Seow

LASALLE College of the Arts, University of the Arts Singapore

Abstract

In this article, I examine aural skills assessment as a chokepoint for decolonizing initiatives in higher music education. While curricular diversification has expanded the range of musical practices encountered in higher education, assessment systems still regulate legitimate listening through narrow epistemic defaults. Rather than offering a comprehensive literature review or proposing a new assessment framework, I advance a theoretical critique of aural assessment as an institutional mechanism for exercising epistemic authority over listening. Using a critical narrative review, I bring decolonizing and intercultural music education scholarship into dialogue with listening epistemologies and assessment theory. I identify recurring failure modes that constrain decolonizing intentions in assessment practice, including construct misalignment, evidence restriction, rubric monoculture, moderation drift, and gaps in examiner preparedness. I then synthesize implications for admissible listening evidence and defensible judgment that preserve rigor without reproducing monocultural standards.

Keywords

Aural skills, assessment, decolonization, listening, epistemology, judgment

© Eugene Seow. The content of this article is the sole responsibility of the author. The ACT Journal and the Mayday Group are not liable for any legal actions that may arise involving the article's content, including, but not limited to, copyright infringement.

Over the past two decades, music education scholars have increasingly addressed decolonization, interculturalism, and epistemic justice, prompting renewed attention to curriculum content, repertoire selection, and the cultural assumptions embedded in pedagogical traditions (Hess 2015; Kallio 2019; Schippers 2010; Volk 2004). Degree programs have diversified stylistically, and genre boundaries have loosened; students are now commonly trained across multiple musical systems, technologies, and modes of practice. However, despite these shifts, assessment practices have often remained comparatively stable, particularly in foundational domains such as aural skills.

In this article, I argue that aural skills assessment constitutes a critical institutional chokepoint for decolonizing initiatives in higher music education. While curricular and pedagogical reforms may signal pluralism in content, assessment practices continue to regulate what counts as legitimate musical knowledge, listening, and competence. In this sense, assessment does not merely measure learning outcomes; it actively shapes what counts as legitimate listening, evidence, and judgment in practice. When assessment criteria remain aligned with narrow epistemic defaults, decolonizing intentions risk being neutralized, as educational values are operationalized in grades, progression decisions, and credentials.

Aural skills occupy a distinctive position within this dynamic. Unlike repertoire-based modules or stylistic studies, aural training is routinely framed as foundational and transferable, underpinning musicianship across genres and contexts. As such, aural assessment carries disproportionate symbolic and practical authority within higher music education. Performance, composition, production, and ensemble work are often interpreted through assumptions established in aural curricula concerning pitch, rhythm, harmony, and musical time. Consequently, the criteria through which aural competence is judged exert a structuring influence on how musical understanding itself is defined and evaluated across programs (Karpinski 2000; Regelski 2016).

It is important to acknowledge that higher music education is not structurally uniform. Aural skills curricula and assessment practices vary considerably across institutions, degree structures, and concentrations, including performance, composition, music education, music technology, jazz studies, therapy, and interdisciplinary programs. In some contexts, aural skills serve as a shared foundational component; in others, listening competencies are embedded within stylistically specific modules or distributed across ensemble and studio practices. The argument advanced here is therefore bounded. It addresses

Seow, Eugene. 2026. Decolonizing aural skills assessment in higher music education: epistemic authority, evidence, and judgment. *Action, Criticism, and Theory for Music Education* 25 (4): 115–38. https://doi.org/10.22176/act25.4.115

contexts in which aural assessment operates as a structurally significant mechanism for progression, credentialing, or benchmarking, particularly where shared assessment frameworks are applied across diverse concentrations. In such settings, treating all students "equally" through uniform assessment criteria may inadvertently obscure the distinct listening demands and musical logics of different practices. A shift from equality toward equity in assessment design—understood as responsiveness to context-specific listening constructs rather than identical evaluative treatment—clarifies the institutional conditions under which aural assessment can function as a chokepoint for decolonizing initiatives.

Existing decolonizing and intercultural music education literature has recognized assessment as a site of power, though it has often received less sustained attention than curriculum design or pedagogical practice (Hess 2015; Human 2019; Tuncer-Boon 2019). When authors discuss assessment, critiques frequently focus on misalignment between diverse learning experiences and monocultural evaluative norms, particularly those grounded in Western tonal–metrical systems and notational literacy. Building on this important work, I focus on how assessment regimes themselves—through constructs, evidence rules, rubrics, moderation practices, and examiner judgment—can reproduce narrow musical ontologies even in ostensibly plural curricula.

Assessment theory provides tools for understanding why such misalignments persist. Foundational and subsequent validity scholarship emphasizes that assessment validity depends not only on task design but also on the coherence of the entire interpretive chain linking construct definition, evidence, and judgment (Messick 1995). In practice-based and performance-oriented disciplines, professional judgment, tacit norms, and institutional calibration processes mediate this chain (Sadler 2009, 2010). When assessment constructs are under-specified or culturally narrow, evaluative decisions may rely on unarticulated standards of "what good sounds like," leading to inconsistencies and inequities that are difficult to identify or contest. Recent scholarship in music education assessment further underscores the need for culturally responsive and context-sensitive evaluative design, particularly in cross-cultural and indigenizing contexts (Human 2019; Human and Akuno 2024; Tuncer-Boon 2019).

In the context of aural skills, this problem is compounded by the tendency to equate listening competence with accuracy-based literacy: rapid identification, symbolic transcription, and alignment within fixed reference systems, including twelve-tone equal temperament and linear meter (Karpinski 2000; Seow 2025). While valuable within many musical traditions, elevating these

Seow, Eugene. 2026. Decolonizing aural skills assessment in higher music education: epistemic authority, evidence, and judgment. *Action, Criticism, and Theory for Music Education* 25 (4): 115–38. https://doi.org/10.22176/act25.4.115

skills to universal benchmarks risks misrecognizing forms of listening developed through oral transmission, embodied participation, cyclic temporal frameworks, or context-dependent intonation (Chernoff 1981; Green 2002). When music educators assess such listening practices through criteria designed for different musical logics, they may interpret differences as a deficit rather than as evidence of alternative competence.

This dynamic produces significant institutional consequences. Aural assessment carries high stakes, influencing progression, degree classification, and access to advanced study (Hess 2015; Sadler 2009). It also plays a role in external examinations and quality assurance, where consistency and defensibility of judgment are paramount. In such contexts, assessment practices tend to prioritize reliability and comparability, often defaulting to familiar constructs and forms of evidence. This institutional conservatism can unintentionally reassert monocultural standards even where curricula and teaching practices have diversified (Hess 2015). As a result, music educators may embrace decolonizing initiatives rhetorically while remaining structurally constrained.

The contribution of this article is analytic rather than prescriptive. Drawing on a critical narrative review, I synthesize literature across (a) decolonizing and intercultural music education, (b) listening epistemologies, and (c) assessment theory on validity, judgment, and standards. I identify recurring assessment failure modes—construct misalignment, evidence restriction, rubric monoculture, moderation drift, and examiner preparedness gaps—and articulate implications for admissible listening evidence and defensible judgment without reproducing monocultural standards.

Defining Decolonizing

In this article, I use the term "decolonizing" in a deliberately bounded and assessment-oriented sense. While decolonization has been mobilized across music education scholarship to address curriculum, repertoire, pedagogy, and institutional culture (e.g., Hess 2015; Kallio 2019), I do not attempt to adjudicate among these broader usages. Instead, I focus on how decolonizing claims encounter—and are often constrained by—formal assessment practices in higher music education. I adopt this usage with awareness of Tuck and Yang's (2012) argument that decolonization is not a metaphor and must entail material restitution and land return rather than symbolic institutional reform. I do not claim that revising assessment practices constitutes decolonization in that material sense. Instead, I adopt a deliberately limited analytic focus on epistemic governance within higher music education: how institutional assessment systems

Seow, Eugene. 2026. Decolonizing aural skills assessment in higher music education: epistemic authority, evidence, and judgment. *Action, Criticism, and Theory for Music Education* 25 (4): 115–38. https://doi.org/10.22176/act25.4.115

stabilize particular listening norms as authoritative. I therefore use the term decolonizing diagnostically rather than declaratively. At the same time, Indigenous scholarship has cautioned that coloniality is enacted not only through territorial dispossession but also through regimes of listening that extract, classify, and adjudicate sound according to settler-colonial epistemologies (Robinson 2020). From this perspective, the governance of listening within institutional assessment can function as a site where such epistemic hierarchies are reproduced, even in the absence of overt curricular exclusion.

Robinson (2020) argues that colonial listening operates through evaluative habits that privilege extraction, categorization, and aesthetic judgment detached from relational accountability to land, community, and sovereignty. In this formulation, listening is not neutral perception but a historically situated practice structured by power. While higher music education assessment is not equivalent to the settler-colonial formations Robinson critiques, the structural logic of legitimating certain sonic interpretations as authoritative—while rendering others peripheral or unintelligible—offers an analytic parallel. Assessment, too, operates by stabilizing which forms of listening are recognizable, defensible, and institutionally valid.

I use decolonizing aural assessment to refer to the interrogation of three interrelated dimensions of assessment practice: epistemic authority, admissible evidence, and ontological defaults. Epistemic authority concerns the listening practices that assessment systems recognize as legitimate and authoritative. Indigenous sound studies further illuminate how colonial habits of perception can structure listening, making Indigenous sonic practices legible only when they are translated into dominant analytic frameworks (Robinson 2020). Such dynamics resonate with assessment contexts that require listening to conform to predefined evidentiary formats before it can count as knowledge. Admissible evidence concerns the forms through which students may demonstrate aural competence. Ontological defaults are the often unspoken assumptions about musical time, pitch, harmony, and structure that underpin assessment criteria and judgments. Assessment norms that presume linear meter, subdivisible time, and resolution-directed harmonic function may misrecognize musics organized around cyclic temporal frameworks and interlocking stratified pulse (as in many Balinese gamelan traditions), or musics in which melodic contour, text, and relational delivery are central to how musical meaning and coherence are recognized (as in many Māori waiata traditions). The point is not to treat these practices as interchangeable, but to illustrate that what counts as stable time, accurate pitch behavior, or coherent form is practice-dependent. Together,

Seow, Eugene. 2026. Decolonizing aural skills assessment in higher music education: epistemic authority, evidence, and judgment. *Action, Criticism, and Theory for Music Education* 25 (4): 115–38. https://doi.org/10.22176/act25.4.115

these dimensions shape how music educators operationalize listening as assessable knowledge.

Black feminist epistemology further clarifies how institutions structure and legitimize epistemic authority. Rather than treating knowledge as neutral or disembodied, Black feminist scholars emphasize the situated validation of lived experience and the politics of recognition that determine whose ways of knowing institutions affirm as credible (Collins 1996, 2022; Nash 2019). Applied to aural assessment, this perspective shifts attention to the evaluative moment itself: which listening practices assessment systems treat as authoritative, whose interpretive frameworks they treat as normative, and how institutional standards naturalize historically dominant sonic logics under the guise of neutrality. This move does not reject standards or expertise; it asks how institutions stabilize standards and whose listening histories those standards reflect.

This bounded use of decolonizing avoids sloganistic or totalizing claims. It does not imply the rejection of standards, expertise, or disciplinary rigor, nor does it advocate relativism in treating all musical responses as equally valid. On the contrary, assessment necessarily involves judgment, comparison, and threshold decisions, particularly in high-stakes higher education contexts. The concern addressed here is not whether judgment should occur, but how the criteria and evidentiary bases for judgment are constituted and normalized within institutional systems.

My analysis focuses on assessment as an institutional practice rather than on classroom instruction. I do not propose a pedagogical program or teaching intervention. I treat assessment as a form of governance that stabilizes particular conceptions of musical knowledge through grading, progression, and credentialing processes. Under this framing, decolonizing assessment entails making visible and contestable the assumptions through which institutions judge listening, rather than prescribing how listening should be taught.

By clarifying the term *decolonizing* in this restricted sense, I aim to reduce conceptual slippage and avoid overclaiming. I do not argue that aural assessment can—or should—resolve all inequities in music education. Instead, I argue that assessment constructs, evidence rules, and judgment practices can constrain decolonizing initiatives precisely when institutions convert educational values into grades, progression decisions, and credentials. In the remainder of the article, I examine how the existing literature illuminates this assessment bottleneck and what implications follow for more defensible and transparent aural evaluation in higher music education.[1]

Seow, Eugene. 2026. Decolonizing aural skills assessment in higher music education: epistemic authority, evidence, and judgment. *Action, Criticism, and Theory for Music Education* 25 (4): 115–38. https://doi.org/10.22176/act25.4.115

Critical Narrative Review

I use critical narrative review here to describe a narrative review organized around theoretical synthesis and critique rather than exhaustive coverage. This approach is appropriate given the heterogeneous nature of the literature examined, which spans music education, assessment theory, and philosophical and sociocultural accounts of listening. These bodies of work differ substantially in epistemological orientation, methods, and evidentiary norms, rendering systematic review approaches neither feasible nor conceptually aligned with the article's aims.

The purpose of the review is not to provide an exhaustive account of all publications addressing decolonization or aural skills, but to synthesize key strands of scholarship that illuminate how assessment practices shape what counts as legitimate musical listening in higher education. I selected sources for their relevance to three intersecting concerns: (a) scholarship on decolonization, interculturalism, and epistemic justice in music education; (b) theoretical and empirical accounts of listening that challenge identification- and notation-centered models; and (c) assessment scholarship addressing construct validity, judgment, standards, and moderation in performance-oriented disciplines.

I use a thematic and interpretive analytic strategy. Rather than aggregating findings, I examine how concepts, assumptions, and evaluative logics recur across literatures, particularly where assessment practices are implicated. I attend to points of convergence and tension between pedagogical aspirations and institutional assessment requirements. Through this process, I identify recurring failure modes that constrain decolonizing intentions in assessment practice, as well as shared implications for evidence, judgment, and standards.

I ground the synthesis that follows in these convergences rather than in the authority of any single theoretical framework. I advance claims cautiously and limit them to what the reviewed literature can reasonably support. When articulating implications for assessment, I frame them as consequences of the synthesis rather than as prescriptive recommendations or evaluative models.

As a narrative review, this approach has clear limits. The analysis is necessarily selective and interpretive, and it does not claim empirical generalizability. However, by drawing together literatures rarely placed in sustained dialogue, I aim to clarify how assessment functions as an institutional mechanism through which epistemic authority is enacted in aural skills education. This methodological stance aligns with the article's diagnostic purpose: to examine why decolonizing initiatives so often stall at the point of assessment, and what existing

Seow, Eugene. 2026. Decolonizing aural skills assessment in higher music education: epistemic authority, evidence, and judgment. *Action, Criticism, and Theory for Music Education* 25 (4): 115–38. https://doi.org/10.22176/act25.4.115

scholarship suggests about more defensible and transparent evaluative practices.

I reconceptualize aspects of aural assessment as epistemic governance: a mechanism that (a) stabilizes ontological defaults about what counts as music, (b) restricts what counts as admissible listening evidence, and (c) disciplines judgment through rubrics, moderation, and auditability (cf. Kane 2006; Messick 1995). This framing helps explain why decolonizing aspirations often succeed at the level of curriculum but stall at the point of institutional decision-making.

Three Literatures as a Problem-Space

Decolonizing and Intercultural Music Education

Scholarship on decolonizing and intercultural music education has expanded substantially over the past two decades, challenging Eurocentric assumptions embedded in curricula, pedagogies, and institutional cultures (Hess 2015; Kallio 2019; Schippers 2010; Volk 2004). Central concerns in this literature include critiques of additive multiculturalism, recognition of epistemic injustice, and the ethical implications of teaching music across cultural boundaries. Hess (2015) argues that simply expanding repertoire or stylistic coverage is insufficient if dominant musical values and hierarchies remain unexamined. Within ACT, decolonization has been examined as a structural problem of disciplinary inheritance and institutional normativity, rather than merely as the inclusion of repertoires (Rosabal-Coto 2019). Recent scholarship has further complicated these discussions by examining how race, gender, whiteness, and institutional pedagogy shape norms within music education, including evaluative cultures (Gould 2021; Grissom-Broughton 2020; Hess 2018; Holder 2025). Such work underscores that decolonizing is not only a matter of curricular expansion but also of interrogating how disciplinary standards are constructed and maintained.

Within this body of work, listening is frequently identified as a critical site of decolonizing practice. Schippers (2010) highlights the need to legitimate diverse ways of hearing, sensing, and engaging with sound, including oral transmission, embodied participation, and culturally specific temporal and pitch frameworks. While these discussions foreground listening as a pedagogical and ethical concern, they often treat assessment implicitly. Authors in this area treat evaluation as consequential, but rarely explain it as an institutional mechanism that enforces epistemic authority.

Seow, Eugene. 2026. Decolonizing aural skills assessment in higher music education: epistemic authority, evidence, and judgment. *Action, Criticism, and Theory for Music Education* 25 (4): 115–38. https://doi.org/10.22176/act25.4.115

When scholars address assessment explicitly, their critiques tend to focus on misalignment between plural pedagogical aims and monocultural evaluative criteria. For example, authors note that students may engage meaningfully with musics organized around cyclic time, non-equal-tempered intonation, or non-functional harmonic logics, yet teachers may assess them using criteria derived from Western tonal–metrical literacy (Hess 2015; Kallio 2019). Such observations underscore the problem but stop short of theorizing how assessment constructs, evidence rules, and judgment practices systematically reproduce these misalignments.

As a result, decolonizing music education literature often identifies what is at stake in assessment without fully articulating how assessment operates as a site of governance. This gap motivates the present analysis, which brings assessment theory into closer dialogue with decolonizing concerns.

Listening Epistemologies

A second relevant body of literature addresses the nature of musical listening itself. Traditional aural skills pedagogy frequently conceptualizes listening as identification and decoding: the accurate recognition, reproduction, and notation of musical elements (Karpinski 2000). Within this model, listening competence aligns closely with symbolic literacy and speeded accuracy, particularly within Western tonal frameworks. Even recent comprehensive treatments of aural skills pedagogy, while expansive in scope, rarely center decolonizing assessment as a structural concern (Cleland and Fleet 2021), reinforcing the need to examine evaluative mechanisms rather than pedagogy alone.

In contrast, scholarship across ethnomusicology, popular music studies, and praxial philosophy emphasizes listening as participatory, relational, and action-oriented (Elliott and Silverman 2015). Authors describe listening as entrainment to shared temporal fields (Clayton, Sager, and Will 2005), embodied engagement in social musicking (Small 1998), and adaptive orientation within culturally specific sound worlds (Chernoff 1981; Green 2002).

Indigenous theorists further argue that colonial listening operates through processes of extraction and evaluation that detach sound from relational, territorial, and community-based contexts (Johnston 2025; Robinson 2020). Such critiques emphasize that listening is embedded within relations of accountability and recognition. When music educators detach sound from these relations and reinterpret it through dominant analytic grids, they transform its meaning. The relevance for assessment lies in recognizing how institutional evaluation may privilege listening practices that align with dominant analytic frameworks

Seow, Eugene. 2026. Decolonizing aural skills assessment in higher music education: epistemic authority, evidence, and judgment. *Action, Criticism, and Theory for Music Education* 25 (4): 115–38. https://doi.org/10.22176/act25.4.115

while marginalizing those grounded in relational or context-bound logics. From this perspective, listening is not primarily about extracting discrete elements but about sensing direction, stability, interaction, and responsiveness over time. Recent accounts of listening as situated, action-oriented sense-making further reinforce the need to treat listening evidence as relational rather than purely identificatory (Rinsema 2017).

These alternative listening epistemologies challenge the assumption that a single set of perceptual skills can function as a universal benchmark for musicianship. They highlight that musical understanding may be demonstrated through timing recovery (regaining shared temporal orientation after disruption), intonational shading (context-sensitive pitch placement within a system), relational adaptation (adjusting one's part in response to others), ensemble coordination, or anticipatory response—capacities not easily captured through notation-centered assessment formats. Importantly, such accounts do not reject rigor; instead, they locate rigor in context-sensitive judgment and relational coherence rather than in symbolic correctness alone.

Despite these insights, scholars rarely translate listening epistemologies directly into assessment constructs. The literature tends to focus on describing how listening operates within particular musical practices, leaving open the question of how such listening might be evidenced, compared, and judged within institutional assessment systems. This unresolved translation problem becomes particularly acute in higher education, where institutions typically standardize, moderate, and defend assessment decisions.

Assessment Theory: Validity, Judgments, and Standards

Assessment scholarship emphasizes that validity is not a property of instruments or tasks alone, but of the interpretive inferences drawn from evidence regarding a defined construct (Messick 1995). From an argument-based perspective, teachers strengthen validity by making the warrants linking observed performances to claims about competence explicit and contestable (Kane 2006). Under this framing, assessment depends on the alignment between what teachers claim to measure, what evidence they treat as admissible, and how they judge that evidence.

A brief terminological distinction is necessary. I use assessment in reference to the processes of eliciting and interpreting evidence about a defined construct of competence, including task design, the specification of admissible evidence, and the inferential warrants linking performances to claims. I use evaluation in reference to the consequential judgment and decision-making

that follows from assessment—grading, ranking, progression decisions, and credentialing. While these processes are often collapsed in practice, distinguishing between them clarifies how music educators exercise epistemic authority: assessment governs what counts as evidence. In contrast, evaluation governs how that evidence is converted into institutional outcomes.

In performance- and practice-oriented disciplines, professional judgment, rather than purely algorithmic scoring, mediates assessment. Sadler (2009, 2010) has shown that such judgment is shaped by tacit standards, exemplars, and disciplinary norms, which may remain under-articulated even as they exert strong influence on outcomes. Although moderation and calibration processes aim to mitigate this subjectivity, they often reinforce shared assumptions about quality rather than interrogating the underlying constructs.

Applied to aural skills assessment, these insights suggest that even well-intentioned curricular reforms may fail to alter evaluative outcomes if assessment constructs remain narrow or implicit. When music educators define listening competence primarily in terms of accuracy-based literacy, they may systemically undervalue alternative forms of listening regardless of pedagogical intent. Moreover, because assessment practices are embedded in institutional quality assurance regimes, they tend to privilege stability, comparability, and defensibility, conditions that can favor familiar evaluative logics over epistemic plurality (Hess 2015).

Taken together, assessment theory highlights that inequities in evaluation are not solely attributable to bias or individual examiners' preferences. They are often produced through structurally normalized practices that render specific forms of evidence legible and others marginal. This perspective is essential to understanding why decolonizing initiatives often stall at the assessment stage, even in institutions that have rhetorically and pedagogically embraced pluralism.

Where Decolonization Fails in Aural Assessment Practice

The literature mapped in the previous section converges on a common tension: while music educators increasingly acknowledge plural listening practices pedagogically, assessment systems in higher music education often remain aligned with narrow evaluative defaults. I synthesize this literature to identify recurring failure modes that constrain decolonizing intentions during aural assessment. These failure modes do not arise primarily from individual bias or resistance,

Seow, Eugene. 2026. Decolonizing aural skills assessment in higher music education: epistemic authority, evidence, and judgment. *Action, Criticism, and Theory for Music Education* 25 (4): 115–38. https://doi.org/10.22176/act25.4.115

but from structurally normalized assessment practices that shape which forms of listening become visible, legible, and comparable.

Construct Misalignment

A primary failure mode concerns misalignment between curricular intent and assessment constructs. In many higher music education contexts, curricula include aims related to intercultural awareness, ensemble responsiveness, or plural musicianship. At the same time, assessment constructs continue to define aural competence primarily in terms of symbolic accuracy and rapid identification (Karpinski 2000). When assessment constructs remain narrow, even well-designed tasks may fail to capture the forms of listening they purport to value.

Assessment theory emphasizes that construct validity depends on coherence between intended learning outcomes and the inferences drawn from assessment evidence (Messick 1995). When aural competence is implicitly equated with notational literacy, assessments may render other forms of listening secondary or invisible. This misalignment allows plural pedagogical aims to coexist with monocultural evaluative outcomes, undermining decolonizing aspirations without overt contradiction.

Exclusion of Admissible Listening Evidence

A second failure mode concerns the restriction of admissible evidence. Aural assessment has traditionally privileged decontextualized demonstrations such as dictation, interval identification, and sight-singing, formats that foreground speed, accuracy, and symbolic translation. Practical tools and strategies for assessing aural skills have been developed to enhance clarity and reliability (Buonviri 2019). However, such approaches still depend on how music educators define constructs and what forms of evidence they deem admissible within institutional systems. While these forms of evidence are appropriate for specific constructs, elevating them to default status limits recognition of listening competencies developed through oral, embodied, or participatory practices (Chernoff 1981; Green 2002).

From an assessment perspective, the problem is not that alternative forms of listening lack rigor, but that institutions often exclude them from evidentiary consideration altogether. Assessment regimes may lack mechanisms to recognize listening demonstrated through timing recovery, intonational shading, or relational adaptation in ensemble contexts. As a result, music educators may assess students with musical expertise grounded in such practices using proxies that do not reflect their actual listening capacities, thereby reinforcing

epistemic hierarchies under the guise of objectivity. From the standpoint of Indigenous listening scholarship, such exclusions may be understood as instances that render relational and context-bound sonic knowledges inadmissible because they do not conform to extractive evidentiary formats (Robinson 2020).

Rubric Monoculture

Even when institutions diversify assessment tasks, rubrics frequently encode monocultural assumptions about musical structure and value. Criteria may appear neutral—emphasizing "accuracy," "clarity," or "control"—while presuming specific ontological defaults, such as linear meter, equal-tempered pitch, or functional harmonic teleology. Black feminist critiques of disciplinary normativity caution that ostensibly neutral standards often encode historically situated power relations, including racialized and gendered assumptions about legitimacy (Collins 2022; Henry 2025). When translated into assessment rubrics, such norms may appear technical while quietly sustaining epistemic hierarchies. Music educators rarely make these defaults explicit, yet they shape the interpretation and scoring of performance and listening.

This rubric monoculture contributes to what Sadler (2009) describes as tacit norming, whereby evaluative judgments rely on shared but unarticulated exemplars of quality. In aural assessment, such exemplars often reflect historically dominant listening practices, making it difficult for alternative musical logics to be recognized as equally coherent or competent. Consequently, rubrics may stabilize narrow evaluative norms even in programs that espouse pluralism.

Moderation Drift and Standard-Setting

Moderation and calibration processes aim to ensure fairness and consistency in assessment. However, in contexts with under-specified assessment constructs, moderation may reassert familiar standards rather than interrogate them. Examiners seeking defensibility and comparability may default to shared auditory expectations—what "sounds right" or "sounds accurate"—without explicitly discussing the underlying assumptions of listening.

This phenomenon, which I refer to as moderation drift, is particularly consequential in aural assessment. Given the ephemeral and context-sensitive nature of listening, moderation often relies on exemplar performances or tacit agreement, rather than on articulated criteria (Sadler 2010). Over time, this can lead to the re-normalization of monocultural standards, even where assessment tasks have been redesigned to accommodate diversity (Hess 2015).

Examiner Training and Epistemic Preparedness

Finally, decolonizing efforts are constrained by gaps in examiner preparation. Examiners are typically selected for their musical expertise rather than for training in plural listening epistemologies or assessment theory. While disciplinary expertise is essential, the absence of a shared language for articulating and evaluating diverse forms of listening can lead to reliance on personal familiarity or preference.

Without explicit support for evaluating listening across multiple musical logics, examiners may struggle to distinguish difference from deficiency. High-stakes contexts, in which accountability pressures incentivize conservative judgment, exacerbate this challenge. The result is not deliberate exclusion but structural inertia: assessment practices continue to privilege the narrow listening norms most readily defensible within existing systems.

Analytic Synthesis: Implications for Aural Assessment

Across the literature reviewed, a consistent pattern emerges: listening practices that do not conform to dominant symbolic and tonal–metrical norms are difficult to recognize, compare, and defend within existing assessment regimes. For aural assessment, this pattern raises two related questions: what forms of listening evidence should count as admissible, and how educators can exercise rigorous judgment without reproducing monocultural standards.

Importantly, the synthesis presented here does not constitute a new assessment framework or evaluative model. Instead, I articulate implications that follow from the reviewed literature, particularly from work on listening epistemologies and assessment validity. I frame these implications conservatively, with attention to institutional feasibility, examiner judgment, and the requirements of moderation and quality assurance.

Admissible Forms of Listening Evidence

A central implication of the literature is that aural assessment requires a broader conception of admissible evidence that recognizes plural forms of listening without abandoning rigor. Across listening epistemologies, scholars have identified several recurring dimensions of aural competence that are not inherently tied to notation or symbolic labeling. These include the abilities to maintain orientation within a shared temporal field (Clayton, Sager, and Will 2005), perceive pitch behavior as context-dependent rather than fixed (Chernoff 1981; Schippers 2010), sense degrees of stability or continuation in

harmonic or textural environments (Huron 2006), and adapt responsively to others in ensemble contexts (Green 2002; Small 1998). From an assessment perspective, these capacities constitute potential evidence of listening, even when demonstrated through embodied, oral, or interactive means.

Rather than treating all such evidence as equivalent, assessment systems should recognize listening evidence that is relational, directional, and context-sensitive (see also Seow forthcoming). This includes evidence demonstrating temporal orientation and recovery, consistency of intonational behavior within a defined system, coherence of responses to perceived musical direction, and responsiveness to ensemble cues. Such evidence aligns with what assessment theorists describe as construct-relevant variance, insofar as it reflects the intended domain of listening competence rather than ancillary skills such as notational fluency alone (Messick 1995).

Crucially, expanding admissible evidence does not entail abandoning standards. Instead, it requires greater specificity about the listening claims being assessed and the forms of evidence through which those claims can be warranted. When collegiate music educators articulate evidence categories more explicitly, they can evaluate plural listening practices for their internal coherence and contextual validity.

Principles for Defensible Judgment

A second implication concerns the exercise of judgment. Assessment scholarship consistently emphasizes that judgment is unavoidable in complex, performance-based domains and that rigor depends on the transparency and defensibility of that judgment rather than on eliminating subjectivity (Sadler 2009, 2010; Seow 2026). In aural assessment, where listening evidence is ephemeral and context-sensitive, judgment plays a particularly central role.

The literature reviewed suggests several principles that support defensible judgment without reverting to monocultural norms. First, assessment requires explicit articulation of the listening frame within which judgment is made. Declaring the musical system, practice, or context under consideration allows music educators to evaluate listening evidence relative to appropriate internal logics rather than against presumed universals (Schippers 2010). Second, collegiate music educators might support listening claims by referencing audible features, rather than asserting them on impressionistic grounds. Whether evidence takes symbolic, oral, or embodied form, teachers might ground their judgments in perceivable musical behavior.

Seow, Eugene. 2026. Decolonizing aural skills assessment in higher music education: epistemic authority, evidence, and judgment. *Action, Criticism, and Theory for Music Education* 25 (4): 115–38. https://doi.org/10.22176/act25.4.115

Third, proportionality between the construct and the criteria is essential. Assessment should not reward or penalize competencies incidental to the listening being evaluated. For example, where the construct concerns temporal orientation or ensemble entrainment, symbolic transcription accuracy may be irrelevant or secondary. Aligning criteria proportionally with constructs supports validity and reduces construct-irrelevant variance (Messick 1995).

Fourth, collegiate music educators could orient moderation and calibration processes toward criteria rather than toward uniform outputs. In plural assessment contexts, music educators can achieve comparability by consistently applying shared evaluative principles to diverse evidence, rather than requiring identical responses. Exemplar discussion and examiner dialogue are therefore central to maintaining standards without suppressing epistemic plurality (Hess 2015).

Finally, defensible judgment depends on examiner preparedness. Where examiners lack a shared language to articulate plural forms of listening evidence, judgment may default to familiarity or personal preference. Examiner development is therefore not an optional add-on but a structural requirement for assessment regimes that seek to recognize diverse forms of listening while remaining accountable (Sadler 2009, 2010; Human 2019; Tuncer-Boon 2019).

Taken together, these principles position judgment as a professional, reflective practice rather than as a neutral technical operation. They suggest that decolonizing aural assessment is less a matter of inventing new tasks than of clarifying how listening is recognized, warranted, and compared within existing institutional structures.

Implications for Institutions, Examining, and Quality Assurance

These implications extend beyond individual aural skills modules. Because assessment practices function as institutional mechanisms of governance, changes in how listening is defined and judged necessarily intersect with program-level decision-making, examining practices, and quality assurance processes.

Implications for Higher Music Education Curricula

Although I do not address curricular reform in this article, it is important to acknowledge the reciprocal relationship between assessment and curriculum. In higher education, assessment practices exert a powerful shaping influence on what is taught, practiced, and valued (Messick 1995; Sadler 2009). Where aural assessment privileges a narrow set of listening competencies, collegiate

music educators may undermine curricular pluralism regardless of stated learning outcomes. Conversely, when assessment criteria recognize a broader range of listening evidence aligned with curricular aims, teaching practices are more likely to reflect that plurality.

From an institutional perspective, this suggests that alignment between curricular intent and assessment construct is a matter of governance rather than of individual teaching preference. Clarifying what constitutes assessable listening does not require expanding syllabi or introducing new content, but it does require coherence between program aims and evaluative practices. Without such coherence, decolonizing aspirations risk remaining aspirational rather than operational.

Implications for Examining and Moderation

Aural assessment frequently involves multiple assessors, external examiners, and moderation processes, particularly in higher music education contexts where comparability and defensibility are paramount. The analysis presented in this article highlights the need for examining practices and moderation processes that can accommodate plural forms of listening evidence without sacrificing standards.

More specifically, examiners might ground agreement in shared criteria and articulated evaluative principles rather than in tacit consensus about correct outputs. Moderation processes that focus exclusively on outcome uniformity may inadvertently suppress epistemic plurality by reasserting familiar listening norms. Scholarship by van der Merwe (2019) on micro-assessment cultures and caring assessment ecologies further suggests that music educators can reorient moderation practices toward shared evaluative principles rather than uniform outputs, particularly in culturally diverse contexts. In contrast, moderation oriented toward construct clarity and evidentiary warrant enables consistent evaluation of diverse listening responses within declared frames of reference. This matters particularly for external examining, where examiners must safeguard standards across diverse cohorts without allowing tacit familiarity to substitute for explicit construct clarity.

Implications for Quality Assurance and Accountability

Quality assurance frameworks in higher education prioritize clarity, consistency, and auditability. From this perspective, plural or intercultural assessment is sometimes viewed as a risk to standards. However, cross-cultural assessment scholarship demonstrates that culturally responsive evaluation can

coexist with standards-based accountability when constructs and criteria are explicitly articulated (Human 2019; Tuncer-Boon 2019). The analysis presented here suggests that such concerns stem less from plurality itself than from under-specified assessment constructs and criteria.

Assessment practices that articulate listening frames, evidence expectations, and judgment principles are more readily defensible within quality assurance systems than those that rely on tacit norms. Explicit criteria do not eliminate professional judgment, but they render it inspectable and accountable. In this sense, attention to assessment design supports, rather than undermines, institutional accountability.

For example, an assessment brief may: (1) explicitly state the listening construct and musical frame presupposed; (2) permit multiple evidence modes aligned to that construct (e.g., embodied reproduction, audio demonstration, relational ensemble response alongside notational representation); and (3) require moderation discussions to reference audible features in relation to the declared construct rather than defaulting to rubric shorthand. These moves do not abandon criteria; they render their epistemic commitments explicit.

Limits and Responsible Use

This argument has several important limits. First, in drawing on Indigenous sound studies, I do not claim equivalence between higher music education assessment and settler-colonial dispossession. Rather, I limit the analytic comparison to the governance of listening as a site where epistemic authority is structured and normalized (Robinson 2020). The argument is therefore deliberately bounded and does not substitute for material or territorial dimensions of decolonization.

Second, readers should not treat the implications articulated in this article as universally applicable across all musical, institutional, or cultural contexts. Higher music education encompasses a wide range of program structures, student populations, and assessment traditions. The assessment failure modes identified here describe recurring patterns rather than uniform conditions, and the relevance of particular implications will vary according to local constraints and priorities. My engagement with Black feminist and Indigenous scholarship in this article is analytic rather than appropriative. I do not intend to collapse distinct historical and political struggles into higher education assessment reform, but to draw on epistemological insights into authority, recognition, and validation to clarify how listening is institutionally adjudicated.

Seow, Eugene. 2026. Decolonizing aural skills assessment in higher music education: epistemic authority, evidence, and judgment. *Action, Criticism, and Theory for Music Education* 25 (4): 115–38. https://doi.org/10.22176/act25.4.115

Third, I do not argue that aural assessment alone can resolve broader inequities in music education. I treat assessment as a critical but partial site of intervention. Structural factors, such as admissions practices, staffing, resource allocation, and institutional histories, also shape which forms of musical knowledge institutions recognize and value. Attention to assessment can mitigate certain forms of epistemic misrecognition, but it cannot substitute for sustained institutional commitment to equity and inclusion.

Fourth, my analysis assumes the continued centrality of professional judgment in aural assessment. While I argue for greater transparency and articulation of evaluative principles, I do not propose eliminating judgment or automating assessment. On the contrary, the defensibility of plural listening assessment depends on examiners' expertise, calibration, and reflexive awareness. Institutions that lack the capacity to develop examiners may find it challenging to implement even modest changes in assessment practice.

Finally, I do not present an assessment framework, rubric, or implementation model. The outlined implications are intended to support critical reflection and dialogue rather than serve as a ready-made solution. I encourage readers to interpret and adapt these implications cautiously, with attention to their own institutional contexts and responsibilities.

Conclusion

I have argued that aural skills assessment constitutes a critical yet underexamined chokepoint for decolonizing initiatives in higher music education. While curricular diversification and pedagogical reform have expanded the range of musical practices encountered in higher education, assessment systems often continue to regulate listening through narrow epistemic defaults. As a result, they may render plural forms of musical understanding invisible or illegible when institutions translate educational values into institutional decisions. The review and synthesis indicate that the problem is structural rather than attitudinal. Decolonizing efforts frequently stall not because of resistance to plurality, but because assessment constructs, evidence rules, and judgment practices remain aligned with familiar evaluative norms. Rubrics, moderation processes, and examiner expectations reinforce these norms by prioritizing stability and defensibility over epistemic transparency.

Addressing this tension requires a shift in how higher music educators conceptualize and govern aural assessment. Rather than focusing solely on curricular content or teaching methods, higher music educators must attend to the criteria through which listening is recognized, warranted, and compared. The

Seow, Eugene. 2026. Decolonizing aural skills assessment in higher music education: epistemic authority, evidence, and judgment. *Action, Criticism, and Theory for Music Education* 25 (4): 115–38. https://doi.org/10.22176/act25.4.115

implications articulated in this article emphasize the importance of clarifying what counts as admissible listening evidence and exercising judgment through explicit, defensible principles that preserve rigor without reproducing monocultural standards (Messick 1995; Sadler 2009). The recurring failure modes identified here reorient attention toward assessment as a site of epistemic governance by clarifying the conditions under which listening becomes admissible, comparable, and defensible within institutional systems. If decolonizing aspirations are to move beyond curricular rhetoric, they must be examined at the point where educational values are translated into evaluative decisions. In this respect, the question is not whether higher music education can sustain plural listening practices pedagogically, but whether its assessment regimes are structured to recognize, warrant, and defend them as legitimate knowledge.

About the Author

Eugene Seow is a Singapore-based music educator and practice-led researcher whose work focuses on aural skills, assessment, musical listening, and epistemic authority in higher music education. He teaches across undergraduate and postgraduate programs in contemporary music and ensemble studies and has served as an examiner and external assessor in multiple institutional contexts. His research examines how assessment structures shape what counts as musical knowledge, evidence, and judgment in plural and intercultural settings. His work has appeared in *Music Education Research*, *Jazz Education in Research and Practice*, *Asia-Pacific Journal for Arts Education*, and the *International Journal of Arts Education*. His broader interests include rhythm pedagogy, ensemble fluency, genre-inclusive curriculum design, and decolonizing approaches to aural practice.

References

Buonviri, Nathan. 2019. Practical tools and strategies for assessing aural skills. In *The Oxford handbook of assessment policy and practice in music education,* volume 2, edited by Timothy S. Brophy, 679–700. Oxford University Press.

Chernoff, John Miller. 1981. *African rhythm and African sensibility: Aesthetics and social action in African musical idioms*. University of Chicago Press.

Cleland, Kent D., and Paul Fleet, eds. 2021. *The Routledge companion to aural skills pedagogy: Before, in, and beyond higher education*. Routledge.

Seow, Eugene. 2026. Decolonizing aural skills assessment in higher music education: epistemic authority, evidence, and judgment. *Action, Criticism, and Theory for Music Education* 25 (4): 115–38. https://doi.org/10.22176/act25.4.115

Clayton, Martin, Rebecca Sager, and Udo Will. 2005. In time with the music: The concept of entrainment and its significance for ethnomusicology. *European Meetings in Ethnomusicology* 11 (ESEM Counterpoint 1): 1–82.

Collins, Patricia Hill. 1996. What's in a name? Womanism, Black feminism, and beyond. *The Black Scholar* 26 (1): 9–17.

Collins, Patricia Hill. 2022. *Black feminist thought: Knowledge, consciousness, and the politics of empowerment*. 30th anniversary ed. Routledge.

Elliott, David J., and Marissa Silverman. 2015. *Music matters: A philosophy of music education*. 2nd ed. Oxford University Press.

Green, Lucy. 2002. *How popular musicians learn: A way ahead for music education*. Ashgate.

Gould, Elizabeth. 2021. Reading Audre Lorde: Black lesbian feminist disidentifications in canonical sociology of music education. In *The Routledge handbook to sociology of music education*, edited by Ruth Wright, Geir Johansen, Panagiotis A. Kanellopoulos, and Patrick Schmidt, 354–65. Routledge. https://doi.org/10.4324/9780429504631

Grissom-Broughton, Paula A. 2020. A matter of race and gender: An examination of an undergraduate music program through the lens of feminist pedagogy and Black feminist pedagogy. *Research Studies in Music Education* 42 (2): 160–76. https://doi.org/10.1177/1321103X19863250

Henry, Jasmine. 2025. Rethinking music theory through Black feminist perspectives. *Journal of Musicological Research* 44 (2–3): 203–14. https://doi.org/10.1080/01411896.2025.2554539

Hess, Juliet. 2015. Decolonizing music education: Moving beyond tokenism. *International Journal of Music Education* 33 (3): 336–47. https://doi.org/10.1177/0255761415581283

Hess, Juliet. 2018. Troubling whiteness: Music education and the "messiness" of equity work. *International Journal of Music Education* 36 (2): 128–44. https://doi.org/10.1177/0255761417703781

Holder, Nathan. 2025. Black feminist thought in UK music education. In *The Oxford handbook of feminism and music education*, edited by Marissa Silverman and Nasim Niknafs, 469–83. Oxford University Press.

Human, Renée I. 2019. Cross-cultural and standards-based assessment of the African musical arts. In *The Oxford handbook of assessment policy and practice in music education,* volume 1, edited by Timothy S. Brophy, 671–96. Oxford University Press.

Seow, Eugene. 2026. Decolonizing aural skills assessment in higher music education: epistemic authority, evidence, and judgment. *Action, Criticism, and Theory for Music Education* 25 (4): 115–38. https://doi.org/10.22176/act25.4.115

Human, Renée I., and Emily A. Akuno. 2024. Indigenising music education: The cross-cultural transfer of African Indigenous concepts and practices. In *Decolonising and indigenising music education*, edited by Te Oti Rakena, Clare Hall, Anita Prest, and Devid Johnson, 88–107. Routledge.

Huron, David. 2006. *Sweet anticipation: Music and the psychology of expectation*. MIT Press.

Johnston, Tory. 2025. *Whales on the mouth: A mixtape of critical Kwinaył sound and music theory*. PhD diss., University of California, Davis.

Kallio, Alexis Anja. 2019. Decolonizing music education research and the (im)possibility of methodological responsibility. *Research Studies in Music Education* 42 (2): 177–91. https://doi.org/10.1177/1321103X19845690

Kane, Michael T. 2006. Validation. In *Educational measurement*, 4th ed., edited by Robert L. Brennan, 17–64. American Council on Education/Praeger.

Karpinski, Gary S. 2000. *Aural skills acquisition: The development of listening, reading, and performing skills in college-level musicians*. Oxford University Press.

Messick, Samuel. 1995. Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist* 50 (9): 741–49. https://doi.org/10.1037/0003-066X.50.9.741

Nash, Jennifer C. 2019. *Black feminism reimagined: After intersectionality*. Duke University Press.

Regelski, Thomas A. 2016. *A brief introduction to a philosophy of music and music education as social praxis*. Routledge.

Rinsema, Rebecca. 2017. *Listening in action: Teaching music in the digital age*. Routledge.

Robinson, Dylan. 2020. *Hungry listening: Resonant theory for Indigenous sound studies*. University of Minnesota Press.

Rosabal-Coto, Guillermo. 2019. The day after music education. *Action, Criticism, and Theory for Music Education* 18 (3): 1–24. https://doi.org/10.22176/act18.3.1

Sadler, D. Royce. 2009. Indeterminacy in the use of preset criteria for assessment and grading. *Assessment and Evaluation in Higher Education* 34 (2): 159–79. https://doi.org/10.1080/02602930801956059

Sadler, D. Royce. 2010. Beyond feedback: Developing student capability in complex appraisal. *Assessment and Evaluation in Higher Education* 35 (5): 535–50. https://doi.org/10.1080/02602930903541015

Schippers, Huib. 2010. *Facing the music: Shaping music education from a global perspective*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195379754.001.0001

Seow, Eugene. 2025. Decolonising the rhythm curriculum: Integrating global groove practices in undergraduate aural skills. *Music Education Research* 27 (5): 593–603. https://doi.org/10.1080/14613808.2025.2565148

Seow, Eugene. Forthcoming. Assessing aural fluency: Toward an anti-colonial, inclusive, and intercultural framework for undergraduate music. *Finnish Journal of Music Education*.

Seow, Eugene. 2026. When formal criteria fail: Judgment and validity in arts assessment. *Journal of Arts and Humanities* 15 (3). https://doi.org/10.18533/zyny3m35

Small, Christopher. 1998. *Musicking: The meanings of performing and listening*. Wesleyan University Press.

Tuck, Eve, and K. Wayne Yang. 2012. Decolonization is not a metaphor. *Decolonization: Indigeneity, Education and Society* 1 (1): 1–40.

Tuncer-Boon, Ebru. 2019. Culturally responsive assessment in music education. In *The Oxford handbook of assessment policy and practice in music education*, volume 2, edited by Timothy S. Brophy, 737–54. Oxford University Press.

van der Merwe, Janelize. 2019. Creating caring micro-assessment cultures in South Africa. In *The Oxford Handbook of Philosophical and Qualitative Assessment in Music Education*, edited by David J. Elliott, Marissa Silverman, and Gary E. McPherson, 431–46. Oxford University Press.

Volk, Terese M. 2004. *Music, education, and multiculturalism: Foundations and principles*. Oxford University Press.

Seow, Eugene. 2026. Decolonizing aural skills assessment in higher music education: epistemic authority, evidence, and judgment. *Action, Criticism, and Theory for Music Education* 25 (4): 115–38. https://doi.org/10.22176/act25.4.115

Note

[1] Generative artificial intelligence tools were used in a limited and transparent manner during the preparation of this manuscript, specifically to support language refinement, structural drafting, and editorial clarity. No AI tools were used to generate sources, analyze literature, verify citations, or produce substantive scholarly claims. All sources were selected and checked by the author, and the author takes full responsibility for the content, argument, and references in this manuscript.

Seow, Eugene. 2026. Decolonizing aural skills assessment in higher music education: epistemic authority, evidence, and judgment. *Action, Criticism, and Theory for Music Education* 25 (4): 115–38. https://doi.org/10.22176/act25.4.115