How you assess your music students may affect how well they learn.

An Emerging Process of Assessment in Music Education

Abstract: An effective assessment process can improve student performance, guide instructional decisions, and advocate for a music program. Strategies include designing and administering reliable and valid measures of student learning and using assessments to enhance feedback, longitudinal documentation of assessment results for accountability, and a transparency of assessment processes and findings for increased advocacy. An emerging process of assessment is inherent through the Model Cornerstone Assessments and contributes to an evolving assessment culture within K–12 music education.

Keywords: assessment, documentation, K–12 music education, Model Cornerstone Assessments

Assessment is an action or instance of making judgment about intended learning, such as a test of knowledge or skill. It’s often considered a “process of gathering and discussing information from multiple and diverse sources in order to develop a deep understanding of what students know, understand, and can do with their knowledge as a result of their educational experiences.” The assessment process can be understood as students demonstrating what they can do and how they apply what they have learned, therefore providing teachers with information that allows for enhancement of instruction. Although not always reflected in practice, school music programs should be designed to develop musical knowledge, skills, competencies, and dispositions sufficient for a lifetime of making music. As a result, music educators can use data collected from assessment processes to share the progress of students with parents, advocate to administrators for resources allotted to music education, and collaborate among faculty to guide curricular decisions for program improvement. The following article defines emerging understandings of assessment

Note: All bolded terms in this article are defined in the Glossary near the end.
and illustrates an assessment process throughout instruction that documents and measures student learning across a range of expectations and informs music instruction.

**An Emerging Paradigm**

In addition to helping students become aware of their progress, assessment is used for improvement and accountability. Schools currently recognize assessment as monitoring and documenting student improvement via evidence gathered. Assessment should illustrate how well students are demonstrating intended learning outcomes. School music teachers use a variety of assessments to identify the extent of student learning. Teachers also use this information to modify pedagogical approaches that could lead to improved student performance or evaluate competency at one point in time. These are often referred to as **formative** or **summative assessments**, respectively. This involves timely, detailed feedback from clearly defined learning expectations and establishes a dialogue between students and teachers regarding the students’ progression toward defined expectations. Feedback can motivate further learning and drive improvement.

Using assessment data for accountability is now an expectation of policymakers and administrators; it is used to demonstrate student learning and teacher effectiveness. Each state uses state or nationally adopted **standards** as a mechanism to provide accountability data in the form of compliance. A limitation, however, is that, as educator Paul Lehman aptly noted, “in the United States, we don’t have an educational system; we have 13,809 educational systems.” The 13,809 school districts across the United States are all varied in both the content and delivery of curricula, offering multiple and diverse opportunities to learn musical content across the spectrum of music students. Therefore, equity and inclusion considerations such as the diversity in appropriateness of curriculum, heterogeneity of student bodies and programs, and opportunity to learn, for example, all challenge universal equity regarding expectations of students’ music learning. The field of music, then, has a philosophical challenge ahead of itself: Does it focus on using a standards-driven approach to music teaching and learning in a way that fosters the normalization and systematization of student learning? Or, does it focus on using a **standards-aligned** approach to music teaching and learning in a way that fosters autonomy of the music teacher or the music student?

Given the importance that public education places on assessment results, music educators should consider aligning instruction and assessment with well-defined standards of learning. Standards-aligned assessments are more effective when embedded throughout instruction at points predetermined by the teacher. Assessment data can enable program improvement through the **documentation** of students’ learning at prescribed points in the learning sequence. For music educators to integrate assessment into instruction, they need to recognize this emerging paradigm of assessment in music education, recognize its value across a music program, and be familiar with a variety of assessment practices that align best with the instructional processes of their schools and districts. Furthermore, these assessment practices should be flexible enough to accommodate differentiation in instruction.

**Defining Assessment in Music**

The public at large often conceptualizes assessment as tests, more specifically, as a system of multiple choice, fill-in-the-blank, or other similar response formats proctored with a specified time limit and in a uniform way, resulting in the conception of the test as the only form of assessment. However, assessment is a process that includes a variety of measures, including a broader definition of test. An assessment process includes
tasks through which students demonstrate proficiencies and achievement of learning outcomes. It also serves as a means of gathering assessment evidence for evaluation and analysis. Expectations are to be based on standards that define what students should know and be able to do reflecting the key concepts, processes, skills, and traditions in music. Reflective of backward design, effective assessment is based on four characteristics: (1) defining the expected learning, (2) determining acceptable evidence of learning by designing tasks that require students to demonstrate necessary skills and cognitive demands, (3) employing a measuring device that differentiates qualities of achievement, and (4) thoughtful analysis and evaluation of the results. An effective assessment process requires collaboration between teachers and students throughout. Assessment discussions reveal areas upon which instructional and curricular adaptations can be implemented to provide a balanced program. Through this process, stakeholders will see the pursuit of instructional improvement and its impact on music programs.

All assessment tasks outline a set of behaviors that reflect intended learning and gather information to draw a conclusion about intended learning based on interpretations about those behaviors. Therefore, the outcome of any test used as an assessment is to make an inference about learning that can be exhibited by a set of observed behaviors. These inferences can be related to the scoring process (e.g., labeling the quality of a performance with a score according to expert opinion or rule); a generalization of the score (i.e., a holistic view of the student’s overall ability of the set of behaviors being evaluated); an extrapolation of a student’s performance into a real-world, authentic context (i.e., vote of confidence of the student’s ability to shift from a controlled testing environment to a noncontrolled authentic environment); or some consequential decision based on the score (e.g., What interventions or processes would the student benefit from based on the results of the test?). Furthermore, when assessment tasks closely reflect how learning occurs in an applied setting, the tasks become authentic assessment; they become relative to a meaningful application of learning in a given context.

To define the quality of generalized inferences about groups of students, it is essential that tests (assessment tasks) be standardized when used as part of assessment processes. Standardization as a process “involves creating conditions that assure uniformity of the tests with regard to administration, difficulty, clarity in the scoring, and establishing psychometric evidence of the quality of the test.” While both are appropriate measures, scoring a student’s musical performance at a solo festival is not equivalent to a student giving a musical performance in the music teacher’s office, thus violating the notion of uniformity in administration. A student giving a musical performance from Book 1 versus a student at the same ability level giving a musical performance from Suzuki Book 2 violates the notion of uniformity in difficulty. A student being evaluated by a lenient adjudicator versus a student of the same ability level being evaluated by a severe adjudicator violates the notion of uniformity in clarity in scoring. From a psychometric perspective, there must be confidence in the inferences made based on the technical properties of the scores. Therefore, the scoring must have evidence of validity and reliability, preferably based on the Standards for Educational and Psychological Testing. Standardization, then, maintains procedural due process of the testing conditions and management of the testing data in the interest of fairness. Fairness, as described in the Standards for Educational and Psychological Testing, can be defined as “[r]esponsiveness to individual characteristics and testing contexts so that test scores will yield valid interpretations for intended uses.” Responsiveness pertains to four important testing categories: (1) test content, (2) test context, (3) test response, and (4) opportunity to learn. Test content addresses whether the content systematically favors or disadvantages some groups over others based on prior knowledge, experiences, level of interest or motivation, or other variables. Test context examines if aspects of the testing environment systematically affect the performance being evaluated. Test response determines whether responses differ based on perceptions of social desirability, background, or any other nonrelated variables. Opportunity to learn focuses on whether the extent to which individuals have had exposure to instruction or knowledge affects their performance.

Regrettably, standardization gets interchanged with the term standard, so it comes with a fair amount apprehension and tension when teachers hear the terms standardize and standard. Standard can have one of several meanings. A standard as a learning goal can be defined as a prescribed level of quality or attainment of a set of knowledge, skills, abilities, or dispositions. Standards can be distinguished into two types: content standards and performance standards. Content standards are “collections of statements that describe specific desired learning outcomes or objectives.” In particular, content standards qualitatively outline the specific knowledge, skills, and abilities a student is expected to know and be able to do at a desired achievement or experience level. Performance standards “specify what level of performance is required for a test taker to be classified into a given performance quality” (e.g., pass/fail, advanced/proficient/basic). These classifications are organized by one or more empirical cut scores that represent a minimum pass level for each respective category. It becomes additionally challenging for music teachers: They often confuse performance standard to mean a content standard describing music performance outcomes. To clarify, music performance (the act of performing repertoire) is actually a content standard, and the level to which students play musically is a performance standard.

Standard-based education, a current trend in schools of adopting standards, continues to dominate the educational landscape in the United States since the publication of the book A Nation
criteria of the new Performance Standards included
culture. Standards were written to
establish expectations in the form of
learning goals within a variety of
disciplinary areas addressed in American
schools. The National Coalition for
Core Arts Standards undertook the task
of reforming the current arts standards
to better establish expectations for the
artistic processes (Create, Perform, Respond, and Connect). As a phila-
osophical tenet, “[t]he central purposes of
education standards are to identify
the learning that we want for all of
our students and to drive improvement
in the system that delivers that learning. Standards . . . to ensure success for
both educators and students in the real
world of the school.”

16 To deviate from
the single mind-set of a standards-based
approach to teaching, there has to be an
alternative to address the variable nature
of the music classroom. Designing tasks
and assessments that are flexible in mul-
tiple contexts are the key to establish-
ing a standards-aligned approach. Task
frameworks that allow for flexibility of
content and skills through which stu-
dents demonstrate learning are designed
and refined. Frameworks of this type
also require scoring devices designed
to accurately measure learning accord-
ing to criteria defined in performance
standards. Effectively designed curricu-
larum and accompanying assessments easily
reflect familiar domains of Bloom’s
taxonomy: 17 the cognitive (knowledge),
psychomotor (skills and executive pro-
cesses), and affective (attitudes, disposi-
tions, and emotional responses). Each
generation of educators, and hence the
standards they develop, exists within a
policy context determined by the needs
and pedagogical preferences of current
culture.

Adding a dimension to the current
standards that had not been addressed
before, the 2014 Music Standards
included Model Cornerstone Assess-
ments (MCAs) that aligned with the
criteria of the new Performance Stan-
dards. MCAs are curriculum-embedded
assessment tasks and measures designed
for music students to apply develop-
mentally appropriate and relevant
knowledge and skills while demonstrat-
ing learning defined throughout the
standards. They are designed to engage
students in tasks authentic to a school’s
curriculum and honor the intent of the
Performance Standards. However, MCAs
cannot only be used to collect assess-
ment evidence, they embody valuable
learning goals and accomplishments
of students within the teachers’ cur-
ricula and can be adjusted to fit any
context. These assessments also pro-
vide a foundation for teachers to collect
and evaluate student work that illus-
trates the nature and quality of student
achievement with increasing sophistica-
tion across grades as envisioned in the

Considerations of curriculum and
assessment should be focused on pre-
paring students for future involvement
with the content beyond their attend-
ce in the school setting or attend-
ance on campus. The MCAs provide a
framework into which teachers integra-
te their own curriculum content to match
the unique goals of any traditional or
specialized music classes while using
common learning expectations of pro-
cesses and rubrics. When validity and
reliability are established, the neces-
sity for external evaluation beyond the
classroom teacher is greatly reduced.
If administered with integrity by prac-
ticing teachers, assessment rubrics like
those provided in the MCAs can reli-
ably illustrate student learning related to
the current music standards as long as
it is clearly understood that additional
assessments of knowledge, technical
proficiencies, and musical skills that
extend beyond the defined standards
must also be administered.

When embedding standard-aligned
assessments within a curriculum, con-
sideration for a school’s culture, context,
curriculum, and opportunities to learn
are essential to ensure appropriateness
and fairness. Respecting the learning
needs of students and the autonomy
of curricula is necessary to allow stu-
dents to demonstrate the quality of
their learning in ways that reflect their
own understandings and competencies.
Effective assessment allows students
to demonstrate musical learning using
a variety of tasks appropriate for their
interests and skills. Among these skills
are the use of technology, composing
new works, improvising, performing
contemporary/pop music, as well as tra-
ditional and nontraditional performance,
informal music-making, and even en-
trepreneurial activities. Each component
(framework, validity, reliability, assess-
ment, scoring, and fairness) is critical
to successfully embedding assessment
throughout a school’s curriculum.

Implementing the Process
An effective assessment process in the
music classroom has six discernable
steps: (1) defining specific outcomes
with expectation levels of achievement;
(2) designing an assessment task and
scoring device; (3) using an enhanced
feedback mechanism; (4) longitudinally
documenting individual student pro-
gress; (5) analyzing the data, improving
practices, and grading; and (6) com-
municating assessment findings to stake-
holders as a means of accountability
and advocacy (see the “Assessment at a
Glance” sidebar).

Defining Specific Outcomes
The first step in implementing an effec-
tive assessment process is defining what
students are to learn and how they will
demonstrate these intended outcomes.
In contrast to goals, which are gen-
eral, broad, often abstract statements
of desired results, outcomes are specific
and measurable and must reflect the
curriculum. They express a benefit or
value added that a student can demon-
strate upon completion of an academic
program or course. In defining learning
outcomes, the challenge often experi-
enced is deciding on specific expecta-
tions that fit into a sequence of learning.
An outcome contains all three of the
following elements: (1) what is to be
learned (knowledge, skill, attitude), (2)
what level of learning is to be achieved

14 A call for rigorous academic
standards emerged from accusations
that public schools demonstrated aca-
demic laxity under a fragmented cur-
iculum. Standards were written to
establish expectations in the form of
learning goals within a variety of
disciplinary areas addressed in American
schools. The National Coalition for
Core Arts Standards undertook the task
of reforming the current arts standards
to better establish expectations for the
artistic processes (Create, Perform, Respond, and Connect).
As a philo-
sophical tenet, “[t]he central purposes of
education standards are to identify
the learning that we want for all of
our students and to drive improvement
in the system that delivers that learning. Standards . . . to ensure success for
both educators and students in the real
world of the school.”
Outcomes are most often defined by developmentally appropriate expectations of the curricular goals, performance standards from state or national standards, or instructor values. They must be translated into language of measurable and observable behaviors.

Assessment Tasks/Scoring Devices

After defining clear learning outcomes and before planning instruction, it is important to confirm how students are to demonstrate the learning defined in the outcome. Assessment tasks should allow students to apply their learning in ways that represent their own understanding and skills in addition to pre-designed expectations of the instructor and program. Accompanying these tasks should be a scoring device to aid in the documenting of student progress during the assessment process. Confirming the reliability and validity of these measures is often difficult for music educators. There are many resources to consult when developing an effective assessment process (see the Resources sidebar); however, all valid and reliable measures consist of carefully designed criteria and levels. Criteria are the specific components of the task that will be observed. Levels specifically define the extent to which students demonstrate each stated criterion. In other words, an effective assessment focuses on what students do and how well they can do them and has the capability to measure qualities of learning.

Feedback

Assessments embedded within instruction are useful to enhance student learning as well as guide instructional improvement. Employing rubrics to assess learning allows teachers the opportunity to provide feedback specific to defined criteria while collecting multiple data points to establish a thorough documentation of student learning. When administered over time, data from scoring rubrics essentially provide a mechanism to provide guided feedback from which students and teachers can monitor growth and learning. Rubrics can simultaneously act as a means of instruction through students being part of developing the scoring criteria and dimensions, student-led feedback sessions, and peer and self-assessment activities integrated into class meetings or rehearsals. These strategies can have a profound impact on the quality of learning that students demonstrate through assessments, and they can be useful to document content learning. Payne found that by using student-created assessments, students approached the tasks with a deeper level of engagement and produced greater specificity in the feedback provided to peers.¹⁹

Documenting

If learning outcomes are well defined and differentiate qualities of learning pertaining to specific criteria, documenting achievement results is straightforward. Many school districts integrate a learning management system (LMS)
that includes an assessment module that maintains assessment scores per student (e.g., Engrade, Thinkwave, Schoology, LearnBoost, Alma with Google Classroom, Canvas, Blackboard, Edmodo, Desire 2 Learn, Moodle). Sometimes these systems have scoring mechanisms, such as program-designed rubrics, that automate the process of documenting and storing achievement scores. When assessments are created to be compatible with the current LMS, data collection and documentation are most efficient. However, vigilance is necessary to ensure that the tasks are authentic to the outcome and not limited by the LMS.

Documentation of student achievement is more than a score from an assessment measure. One way to document progress beyond the score is through audio or video recordings. With devices becoming increasingly portable, smartphone and tablet technologies are useful in documenting student work. Students can record themselves, complete a self-assessment (or peer assessment), and then upload the documents into an LMS or on-site storage system maintaining a record that documents each developmental level of progress. Archiving of student work and outcomes assessment scoring provides a powerful means to longitudinally substantiate progress or expose learning needs for both the teacher and the student as well as establish accountability within the program. Examples of student work can also be used as exemplars for illustrating student achievement. Teachers can inform future instructional practices using multiple data point information through the application of rubrics, checklists, tests, self- and peer assessments, scored assignments, or alternative opportunities to demonstrate student learning and guide further learning.

Analysis

Thoughtful analysis of the assessment data is one of the most valuable components of an assessment process. Assessment scores from summative assessments are to be reviewed to identify student achievement and effectiveness of the current curricular sequence, expose learning needs of students, and guide instructional choices. Individual student scores over time provide a picture of student learning achieved and areas of learning not yet attained. One misconception is that the term grading can, or should, be used interchangeably with assessment. These terms are not synonymous. Assessment is the measurement and evaluation of a student’s performance, whereas grading is the assignment of value (e.g., grade) to that performance or growth demonstrated. Growth is a value assigned by the teacher considering the (1) difference in assessment score from one point to the next, (2) difficulty of the task, (3) initial level of proficiency, (4) learning aptitude of the student, and (5) intended expectation for student in the program context. Value, within the context of student learning, is seldom reflected in a score from a singular assessment rather than a set of multiple, thoughtfully designed and sequenced assessments. Analysis is aided by reliable, valid, and fair scoring devices. Ultimately, grading practices are up to the teacher and should document and reflect student growth. Employing an effective assessment process will provide indicators through continuous data collection to accurately make determinations.

Communication and Advocacy

One benefit of a thorough assessment process is the establishment of a framework of accountability that promotes and supports the quality of a music program to a variety of stakeholders such as students, parents, administration, and the extended educational community. Transparency enables these stakeholders to make informed judgments through which enhanced support can be garnered. However, not all assessment data are intended for external stakeholders; some assessment data can easily be misunderstood in a broader context beyond a classroom or program. But when an assessment process can document qualities of learning in valued outcomes and student learning is responsibly reported, perceptions of educational value and rigor of programmatic quality can be greatly enhanced. Essentially, the assessment process can provide critical and invaluable information to advocate for the existence or expansion of a music program. Teachers can use findings derived from analyses to provide a strong foundation for accountability in the classroom and justify effective teaching practices. While classroom assessments are not a measure of teaching effectiveness, they can be a strong indicator of successful practices.

This emerging process of assessment is inherent through the recent release of the Model Cornerstone Assessments and contributes to an evolving assessment culture within K–12 music education. Effective assessment practices can improve student performance, guide instructional decisions, and advocate for a music program. Assessing student learning is one of the primary responsibilities of a music educator, an essential component of instruction, and a process that must be embraced by the profession. Understanding of and competency with effective assessment processes ensures student learning at all levels and creates strong and dynamic music programs. Addressing these issues must be accomplished through both professional development opportunities and inclusion in teacher education programs. As a profession, music educators must adapt to this emerging assessment process to allow for increased student learning and greater transparency for all involved.

Notes

2. Mary E. Huba and Jann E. Freed, Learner Centered Assessment on College Campuses: Shifting the Focus from Teaching to Learning (Boston, MA: Allyn and Bacon, 2000), 8.
Assessment integrated into instruction can inform both the teacher and the music learner.

(Article continues with Glossary on pp. 43 and 44.)
### GLOSSARY

- **Assessment (Process)** – The collection, analysis, interpretation, and applied response to information about student performance or program effectiveness in order to make educational decisions resulting in continual improvement.
- **Assessment Task** – Activity or test designed to measure and evaluate a student’s demonstration of attained knowledge, specific skills, or expected dispositions.
- **Authentic Assessment** – An assessment task that reflects the way a student thinks and interacts in an environment connecting with belief and experiences within each individual student.
- **Benchmark** – a standard or point of reference against which things may be compared or evaluated.
- **Content Standards** - collections of statements that describe specific desired learning outcomes or objectives.
- **Criterion** – a defined expectation of the standard of performance for a particular task. (plural: criteria)
- **Dispositions** – a person’s inherent or developed qualities of mind and character.
- **Documentation** – The method used to record student achievement through scores and illustrative examples for evaluation and accountability.
- **Evaluation** – The collection and use of assessment data to make judgments of student achievement or program effectiveness to make informed educational decisions.
- **Exemplar** – a person or thing serving as an excellent model.
- **Fairness** – Responsiveness to individual characteristics and testing contexts so that test scores will yield valid interpretations for intended uses.
- **Formative Assessment** – Ongoing assessment within an educational program for the purpose of exposing learning needs and guiding improvements.
- **Grading** – The assignment of value to that performance or growth demonstrated.
- **Illustrative Example** – Student work that serves as a model for meeting a specific set of criteria.
- **Learning Management System (LMS)** - a software application for the administration, documentation, tracking, reporting and delivery of educational courses or training programs.
- **Level** – A descriptor that describes the knowledge, skill, and/or disposition performance level of students to allow teachers to determine an achievement score.
- **Measurement** – The use of systematic methodology to observe musical behaviors in order to represent the magnitude of performance capability, task completion, and concept attainment.
- **Model Cornerstone Assessments** – curriculum-embedded assessment tasks and measures designed for music students to apply developmentally appropriate and relevant knowledge and skills while demonstrating learning defined in standards.
- **Performance Assessment** - An assessment requires students to demonstrate learning that requires integration of expected knowledge, skills, and/or dispositions.
- **Performance standards** – Described level of performance required for a test taker to be classified into a given performance quality.
- **Portfolio Assessment** – An analysis of a collection of student work used to demonstrate student achievement in a content area; student progress is determined by reviewing the collected works in light of previously established criteria.
- **Reliability** – The consistency of an assessment task or tool to produce similar results over a given amount of time.
- **Responsiveness** – the ability of a measure to change over a pre-specified time frame and the extent to which change in a measure relates to corresponding change in a reference.
GLOSSARY (continued)

- **Rubric** – A set of scoring criteria used to determine the value of a student’s performance on assigned tasks; the criteria are written so students are able to learn what must be done to improve their performances in the future.
- **Self-Assessment** – Analysis of one’s own knowledge, skills, and dispositions.
- **Standard** – The content, level, or type of performance expected of students at a particular point in time or stage of development.
- **Standardize** – To cause an assessment to conform to a standard.
- **Standards-Aligned** – A system of instruction, assessment, and documentation focused on students’ demonstration of knowledge, skills, and dispositions to a mutually agreed upon set of intended student learning outcomes or expectations as they progress through their education.
- **Summative Assessment** – An assessment at the end of an instruction cycle to measure student growth and learning.
- **Validity** – The effectiveness of an assessment instrument in measuring what it is supposed to measure; also the appropriate use of assessment data in reporting and analysis.

* All definitions were synthesized from multiple sources by the authors.