



Title: Using Planned Assessment Techniques to Increase Student Growth in Music Education

Author(s): George L. Duerksen

Source: Duerksen, G. L. (1995, Winter). Using planned assessment techniques to increase student growth in music education. *The Quarterly*, 6(4), pp. 57-64. (Reprinted with permission in *Visions of Research in Music Education*, 16(6), Autumn, 2010). Retrieved from <http://www-usr.rider.edu/~vrme/>

Visions of Research in Music Education is a fully refereed critical journal appearing exclusively on the Internet. Its publication is offered as a public service to the profession by the New Jersey Music Educators Association, the state affiliate of MENC: The National Association for Music Education. The publication of VRME is made possible through the facilities of Westminster Choir College of Rider University Princeton, New Jersey. Frank Abrahams is the senior editor. Jason D. Vodicka is editor of the *Quarterly* historical reprint series. Chad Keilman is the production coordinator. The *Quarterly Journal of Music Teaching and Learning* is reprinted with permission of Richard Colwell, who was senior consulting editor of the original series.

Using Planned Assessment Techniques To Increase Student Growth In Music Education

By George L. Duerksen

University of Kansas

Music education in the United States has a long, strong, on-going tradition of assessing musical performance. Music educators traditionally have been accountable on the basis of public assessments. The crowd's judgment of the band show at half-time, the judge's rating earned by the choir at a festival, the service club's evaluation of the noon-time performance by the showchoir, and the parents' and teachers' satisfaction with the music presented at a PTA meeting regularly provide assessments concerning the music education program. These assessments concern specific, isolated performances, however, without careful attention to the relationship of those performances to student learning.

Current school improvement models, when adapted specifically to school music programs, suggest valuable ways to improve stu-

Current school improvement models, when adapted specifically to school music programs, suggest valuable ways to improve student progress toward the goals which the music education profession values most highly.

dent progress toward the goals which the music education profession values most highly. Regional accrediting associations such as the North Central Association of Colleges and Schools (NCA) and some state education agencies require that data-based school improvement programs be in operation. These school improvement models place strong emphasis on using assessment data systematically throughout the educational process. The plans typically require ongoing, multiple assessments of student growth and program effectiveness. While the individual components of these assessment programs are not new, their combination as used in the school improvement and accreditation context might prove to be a helpful model for music education.

Characteristics and Organization of the School Improvement Models

School improvement models clearly identify the target goals of the school's educational program, and then use frequent, regular, and multi-faceted assessments to document student progress toward those objectives. The assessments also provide informa-

George L. Duerksen is Professor and Director of the Music Education/Music Therapy Division at the University of Kansas. Research interests include: music psychology, functional music, and measurement techniques.

Data from multiple sources are used to make holistic and analytic judgments concerning student progress and achievement as well as the school's instructional effectiveness.

tion for designing specific curricular and instructional processes designed to help students reach goals and to monitor and adjust the instructional processes when needed.

The models include:

- orientation toward the future.
- focus on student growth, using comparable measures before, throughout, and at the end of the instructional process to assess this growth.
- pervasive and systematic use of multifaceted assessment data throughout the planning and instructional process.
- extensive use of both holistic and analytical assessments, using measures appropriate for specific purposes.
- priority on application and assessment of knowledge and skills in real-life situations that require higher level thinking skills.
- precise alignment of (a) instructional goals/objectives; (b) instructional/learning activities; and (c) assessments/analyses.

These models provide a continuous attempt to improve instruction and student achievement. Each iterative cycle attempts to improve on the prior work.

The processes described in this article derive primarily from two accreditation models. They are the Outcomes Accreditation endorsement used by the Commission of Schools of the NCA and the Quality Performance Accreditation process of the Kansas State Board of Education (North Central Association, 1994, Kansas State Board of Education, 1993). In these models, assessment data play a broad and continuous role, pervading the entire process. Data from multiple sources are used to make holistic and analytic judgments concerning student progress and achievement as well as the school's instructional effectiveness. The assessment data provide:

- information used in goal selection.

- information concerning student's initial status in relation to the goals.
- suggestions for the curriculum and instruction most likely to help students progress toward the goals.
- information concerning student's progress toward the goals.
- information for monitoring and adjusting curriculum and instruction on the basis of ongoing student progress.
- information documenting student's achievement of the goals.
- information used in updating goal selection for each cycle of the continuous improvement process.

The NCA model places particular emphasis on precise alignment of goals, outcomes, curriculum, instructional interventions, and assessments — in other words, coherence among the various parts. The school community establishes target goals that reflect clearly what students should know and be able to do when they leave the school's program. Curriculum, instructional strategies, and assessments are then aligned with those goals.

The school improvement process begins by examining the existing program and its context, current student knowledge and skill levels, and the needs the students will likely face in the future. The school community uses data from this assessment to support selection of student outcome goals.

School improvement plans often emphasize goals that cross traditional curricular areas. The goals often focus on higher order thinking skills, such as those described in the *Taxonomies of Educational Objectives* (Bloom 1956, Krathwahl, Bloom, Masia 1964). Goal areas often include communications, problem solving, caring for self and others, and skill in working individually and collaboratively. Interdisciplinary approaches as well as instruction in traditional disciplines are used to help students grow toward these goals. Assessments in interdisciplinary set-

The NCA model places particular emphasis on precise alignment of goals, outcomes, curriculum, instructional interventions, and assessments — in other words, coherence among the various parts.

tings and the traditional disciplines are used to monitor progress toward these goals. Such outcomes can be important in the music program; they can be pursued and assessed in the music context.

Once the student goals have been specified, the school community develops an action plan. This plan consists of strong interventions (instructional strategies) and assessments to use in pursuit of those goals. Often a broad array of curricular and extracurricular influences within the school program focus on the specific goals. The assessments, likewise, are designed for administration across the curriculum in a variety of settings.

The assessment plan provides for multiple measures, across time, in an ongoing system. Repeated measures, focusing on those complex goals that are deemed most important, are organized to track student growth. This “early-during-late in the process” assessment design provides information about student progress and achievement, and the effectiveness of instructional strategies as well.

The action plan timeline is sufficiently long — often a number of years — that there can be reasonable expectation of achieving substantial student progress toward the goals. During implementation, ongoing assessment data are used to modify the course of action as appropriate. At the end of the implementation period, student knowledge and skills in the target areas are again assessed. The school accreditation model then uses the wide variety of assessment data to make a holistic judgment of a school’s success in achieving student progress during the improvement cycle. Results of this final assessment are used as part of the initial step of the next iteration of the improvement cycle. This model might be applied to judge student progress toward the final objectives of the music education curriculum. The model depends on assessment and the thoughtful

use of the resulting data.

The accreditation process suggests that each discipline, such as music, has identified what students should know and be able to do when they exit from that field’s curriculum. Again, multiple assessments within each curriculum area are expected to be aligned with that discipline’s over-riding goals.

The Assessment Plan

The school improvement approach suggests using both “holistic” and “analytic” techniques in the assessment plan. Holistic assessments attempt to describe the “big picture” of complex behaviors. Analytic assessments provide more specific data concerning detailed, individual behaviors. In either case, the following questions underlie selection of measures:

- What indicators will be sufficient to demonstrate that a student has reached those outcomes?
- What indicators will be sufficient to demonstrate student progress toward those outcomes?
- What indicators will be sufficient to indicate whether the curriculum is effectively helping students reach those outcomes?

The selected assessments are expected to meet a variety of purposes.

Parents and policy makers want relevant assessments that reflect the capabilities our children will need to become successful adults in a complex society. Educators want assessments that function as an integrated part of the instructional process; assessments that model high quality instruction and promote student learning. (ETS, 1995)

Clearly, these purposes require multiple, ongoing, assessment of student achievement in a variety of challenging situations. Since assessment is not a goal in itself, but a tool to support student learning, it needs to teach, be unobtrusive, and focus attention and time on instructional goals.

A useful categorization is made by Wick

Judging and interpreting student performance in complex on-demand and context-bound measures requires special techniques.... In many cases a scoring rubric is developed to use in judging the student's achievement in task performance.

and Gose (1994) when they describe assessments as being "common metrics" or "context bound." Common metrics are measures which, if applied appropriately, allow comparisons among and between students, schools, districts, states, or nations. Common metric measures, such as standardized achievement tests, state-developed achievement tests, and the like, provide information of interest to policy-makers such as legislatures, state education agencies, school boards, and the general public. These measures are also of some interest to students and parents, especially in the later stages of the student's academic career.

Context-bound measures, on the other hand, are those that are specific to individual classrooms and schools. They are of primary value in the school setting. They tell clearly what students know and can do, and how instruction is going (Wick and Gose, 1994). They capture what actually goes on in the classroom and with individual students. These measures provide the information for adjustments in curriculum and instruction, and data about the growth of individual students. Context-bound measures provide data that are of particular interest and use to students, parents, and teachers.

Common metric measures can be divided into two groups: standardized tests, and "on demand" performance measures. Standardized tests are sometimes described as "performance indicators" because their results are interpreted as indicating what a student might do in a real performance situation. On-demand performance measures ask all students to engage in a complex, authentic performance task in an appropriate context. In the NCA model, "on demand" assessments ask all students to engage in the same task at specified times during the instructional year, and in some cases from year to year. Thus the

results of these measures can show student improvement in these tasks during the ongoing course of the action plan's operation.

Context bound measures typically can be categorized as either "curriculum embedded" or "portfolio/artifacts" (NCA, 1994). "Curriculum embedded" simply refers to measurements characteristically used in the course of instruction. In textbook-based classes these might include tests at the ends of chapters, mid-term and term-end measurements, and the like. Portfolio/artifacts serves as a generic phrase for an accumulation of student work. Portfolios vary. They often collect examples of a student's "best work" at various times through the course of instruction. They may also include students' reflections on their own work and progress. In any case, the information in a portfolio allows an assessment of student change over time.

In the assessment plan design, a strong link needs to exist between the context bound and common metric assessments (Wick and Gose, 1994). Since the common metrics are particularly important to policy-makers and the general public, students need to show progress in those measures. Focus on specific expected items of standardized measures, or expected details of authentic, complex, performance tasks — teaching to the test — constricts the curriculum and limits the educational value of instruction. In contrast, context bound assessment can have the breadth and variety needed to represent the full universe of knowledge and skills deemed to be important. Authentic instruction and assessments embedded in the daily educational process can encourage and monitor student progress. Pupils can learn and practice skills directed toward selected goals, and in the process demonstrate those skills in complex, authentic situations. Here, "teaching to the test" can be a virtue if the

Reliability and validity of authentic assessments are important concerns.

assessments are multiple, ongoing, and comprehensive in reflecting the total of the instructional objectives. Then, if the context bound assessments and common metrics are appropriately aligned, student growth in the former should result in increased scores on the latter.

On-demand and context-bound assessments can be designed to assess student behavior in complex, authentic situations that challenge higher level thinking skills. Wiggins (1991) lists three principles common in design ideas for authentic performance tasks:

- 1) higher-order thinking and acting require that students produce “unique products or performances” (in the words of Bloom et al.)
- 2) assessment tasks should teach; the best ‘tests’ educate students about the types of challenges actually encountered ‘in the field’ when professionals are called upon to *use* knowledge effectively and imaginatively.
- 3) authentic tasks require the kinds of judgments that are routinely called for in real-world ‘tests’ of knowledge. The best ‘answers’ are sensitive to the context of the situation, such as: ‘audience’, particular constraints and purpose of the setting, appropriate precision, and cost/benefit considerations.

In this sensitivity to context, authentic measures differ substantially from the traditional approach to test development. “Unlike objective tests, which purposefully attempt to remove context from the items, performance assessment tasks intentionally capture performance in context, and the response to the task is determined by factors that go much beyond the structure of the task itself” (Delandshere and Petrosky, 1994, p 9).

Judging and interpreting student performance in complex on-demand and context-bound measures requires special techniques. The authentic assessment process has several parts:

- a stated outcome or goal;
- a specified performance;
- a situation, prompt, or exercise that elicits the performance; and
- specific evaluation criteria for the performance (Petz, 1995).

In the instruction and assessment process, students are made aware of the details of the assessment methodology. Students learn the intended goals, the performance expectations, the types of exercises, and the specific evaluation criteria that will be used. As Petz points out, the emphasis is on “mastery learning” rather than “mystery learning.”

In many cases a scoring rubric is developed to use in judging the student’s achievement in task performance (Mid-Continent Regional Educational Laboratory, 1992). The rubric consists of two parts. One part is a scale, using words or numerals. This scale usually indicates different degrees of quantity. The other part is a set of descriptors associated with particular locations on the scale.

To develop the descriptors, the components of the successful performance need to be described. Levels of accomplishment of these components are then assigned to the particular locations of the scale. Useful accessories include example performances that illustrate each value on the scale. These “performance anchors” provide benchmarks with which to compare student performances and prove useful in training individuals to use the rubrics.

Holistic and analytic rubrics are both useful. In the holistic mode, a single overall score is assigned to the performance task. This is similar to the judges’ assessment of a performance at contest, except that the descriptors may provide a more specific understanding of the “meaning” of that score. Holistic descriptors help the rater assess the overall performance — the total effect of all of the specific components and those not identified as they combine in the actual performance. In the analytical mode, several

Specific program improvement plans for music education might begin with examination of the profession's most valued goals, as adapted to individual schools.

parts of a task are scored separately. Analytic rubrics focus on specific components of complex performance tasks, and rate student performance on each one. For instance, an analytic rubric for a music performance might give separate ratings for intonation, tone quality, rhythmic accuracy, note accuracy, articulation, dynamics, and other components.

Holistic rubrics are useful for tracking student improvement over time, and documenting the overall effectiveness of interventions in a program improvement process. Analytic rubrics can be particularly helpful in choosing instructional goals for particular students. They allow identification of specific strengths and weaknesses. Students can use the information for setting and pursuing individual goals, and teachers can use the information to choose specific instructional emphases for each student.

Reliability and validity of authentic assessments are important concerns. LeMahieu, Gitomer, and Eresh (1995) provide a thoughtful analysis based on their review of a variety of studies that show varying degrees of stability in rater's judgments and their own study of district-wide portfolio assessment in the Pittsburg schools. Apparently, extensive involvement in development and interpretation of the scoring rubric, leading to shared understanding of the rubric by raters, is an important factor in increasing inter-rater reliability. However, "In a classroom context, where one teacher evaluates all of the student solutions or products, the scoring criteria are likely to be consistently applied" (ETS Trustees' Colloquy, 1995). The school improvement model uses multiple assessments in varying contexts to help assure the accuracy of overall judgment of individual student growth. In the final analysis, as Wick and Gose (1994) indicate, "since measurement error can never be eliminated, consumers should be knowledgeable about measurement error in making inferences and

taking actions" (Wick and Gose, 1994, p.165).

Validity concerns are addressed in several ways. The development of assessments that are carefully aligned with instructional interventions and the instruction's target outcomes should help assure construct validity. Messick (1989) suggests that the adequacy and appropriateness of inferences and actions based on assessment results defines validity. This view, again, supports the importance of the alignment of goals and assessments.

Music Education Applications of the Models' Characteristics

Specific program improvement plans for music education might begin with examination of the profession's most valued goals, as adapted to individual schools. The Music Educators National Conference (1986) lists ten intended outcomes for the music education program. The National Standards for Arts Education (MENC 1994) and *Arts Education Assessment Framework* (National Assessment Governing Board 1994) provide excellent suggestions for potential goals and assessments that might be adopted for their particular school community contexts.

To parallel the school improvement model, the goals suggested in these sources need to be examined in the context of the challenges the school's particular students are expected to face in the future as well as the values of the school community. What music-related knowledge and skills will students need to cope with anticipated and unanticipated future challenges? What music-related skills and knowledge are valued most highly? Careful gathering and analysis of assessment data can provide information to help answer these questions. This assessment most frequently examines the needs and aspirations of students, parents, teachers, and community (broadly conceived), as well as needs for the future. Analysis of these data provides the foundation for deciding on the program goals.

A complex picture of the “musically educated person” may appear in the program goals. . . . A variety of knowledge, musical skills, and dispositions characterize this person. Multiple assessments of a variety of types are most likely to support accurate evaluation of progress toward such complex goals.

Once the music target goals have been chosen, assessment plans need to be aligned with them. A complex picture of the “musically educated person” may appear in the program goals. The musically educated person integrates and demonstrates knowledge, skills, and dispositions in ways that differ from the simple sum of individual musical behaviors. A variety of knowledge, musical skills, and dispositions characterize this person. Multiple assessments of a variety of types are most likely to support accurate evaluation of progress toward such complex goals. Currently available standardized tests in music do not seem sufficient to meet the task. Neither will data from measurement of small-scale musical tasks accumulate sufficiently to assess progress toward or achievement of the complex goal.

The music assessment plan identifies common metrics and context-bound measures that would provide holistic and analytic evaluation of student growth toward the valued goals. Some of these measures might be readily available. There may be standardized tests or currently-used classroom measures that align with the selected goals. The music educator will probably need to develop new authentic assessments that fit directly with the selected program goals. Some of these will be on-demand assessments that serve as common metrics; others will be context bound measures that will be integrated into the day-to-day instructional process.

The assessment plan may include a portfolio that collects examples of “best work” through the course of instruction and students’ reflections on their own work and progress (Goolsby, 1994, 1995). In any case, the information in the portfolio allows an assessment of student change over time. In a

music class, the portfolio might contain recorded samples of a student’s playing, singing, or compositions taken at different times during the instructional year. The plan will also identify a holistic assessment process to be used at the end of the improvement cycle to estimate the program’s overall success in helping students reach the target goals.

The plan’s assessment timeline will schedule the various measures over the full improvement cycle. The repeated measures in this timeline allow assessment of student growth as the cycle proceeds.

Having such an improvement plan provides a way for music education to help focus student and teacher attention on musical knowledge and skills valued most highly while paying appropriate attention to specific details and public performance. In addition, music educators can develop ways to assess, in the context of the music program, student growth toward the school-wide, cross-curricular goals selected in specific school improvement plans. Ultimately, a well-developed assessment plan can help the music program play an integral role in the school improvement process, in assuring student progress, and in designing curriculum and instruction.

References

- Bloom, Benjamin S. (ed.). (1956). *Taxonomy of educational objectives, handbook II: Cognitive domain*. New York: David McKay.
- Delandshere, G. & Petrosky, A. (1994). *Do the fundamental assumptions of educational measurement hold for the assessment of complex performances?* Draft paper presented at the annual meeting of the American Educational Research Association, Session 19.65, New Orleans, April 1994.
- Educational Testing Service. (1995). *Trustees’ Colloquy — Performance assessment: Different*

- needs difficult answers.* Princeton, NJ.: Author.
- Goolsby, T. W. (1994). Music portfolios in secondary music classes: What do they measure? *Special Research Interest Group in Measurement and Evaluation*. 17, 11-18.
- Goolsby, T. W. (1995). Portfolio assessment for better evaluation. *Music Educators Journal*. 82(3), 39-44.
- Kansas State Board of Education. (1994). *Assessment! assessment! assessment!*. Topeka, KS: Author.
- Krathwal, D. R., Bloom, B. S., and Masia, B. B. (1964). *Taxonomy of educational objectives, handbook II: Affective domain*. New York: David McKay.
- LeMahieu, P. G., Gitomer, D. H. & Eresh, J. T. (1995). *Portfolios beyond the classroom: Data quality and qualities*. Princeton, NJ: Educational Testing Service.
- Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational Measurement*. New York: Collier Macmillan.
- Mid-Continent Regional Educational Laboratory. (1992). *Guidelines for writing rubrics*. Aurora, CO: Author.
- Music Educators National Conference. (1994). *National standards for arts education*. Reston, VA: Author.
- Music Educators National Conference (1986). *The school music program: Description and standards (2nd ed.)*. Reston, VA: Author.
- National Assessment Governing Board. (1994). *Arts education assessment framework*. Washington, D.C.: Author.
- North Central Association Commission on Schools. (1994). *A focus on student success*. Tempe, AZ: Author.
- Petz, Jean Ann. (1995). *Interventions, assessments, and documenting student success*. Wichita, KS: Kansas North Central Association.
- Wick, J. & Gose, K. (1994). *Improving student performance in your school*. Dubuque, IA: Kendall/Hunt.
- Wiggins, G. (1991). *"Task" design ideas, principles and guidelines*. Geneseo, NY: Center on Learning, Assessment, and School Structure.



UNIVERSITY OF NORTHERN COLORADO

SCHOOL OF MUSIC

- Nationally acclaimed faculty, with specialists on every instrument
 - More than 50 award-winning performance ensembles
- Full range of degrees from bachelor through doctoral programs:
BA, BM, BME, MM, MME, DA, DME
- Emphases in music education, music history/literature, theory/composition, music performance, conducting, performance/pedagogy, music theatre, opera theatre, jazz studies
- Employment opportunities, scholarships and assistantships for qualified applicants

Call (970) 351-2678 or write:
University of Northern Colorado
College of Performing and Visual Arts
School of Music
Greeley, CO 80639

The University of Northern Colorado is fully committed to Affirmative Action and Equal Opportunity