Are you at a loss for ways to assess your music students? Mitchell Robinson offers some ideas that might help you improve your situation.

BY MITCHELL ROBINSON

Phil came home well after 6:00 P.M. feeling more tired and frustrated than he had all year. There had been the conversation with the principal, the three angry phone calls from parents of kids in his band, and then, at the very end of the day, his best trumpet player and set drummer told him they were dropping his class from their second-quarter schedules. All of this occurred as a direct result of the grades Phil had posted that morning after being up until 2:00 A.M. hastily “bubbling-in” numerical grades on the computer-generated forms that the school had adopted only the previous semester.

Phil knew he should have kept better records on his students during the term. One of his recurring nightmares involved being called on the carpet over a poor grade and not having any tangible evidence to justify the mark he gave. He always felt a twinge of guilt at report-card time as he eyeballed his students’ names: “I think an 85 for Tim should do; he does a pretty good job on the bass clarinet and never gives me a hard time. Renee is a different story; she hasn’t come to a lesson in three weeks! I’ll give her a 60 this marking period—maybe that’ll shake her up!” Phil knew there had to be a better way, but was at a loss for ideas.

Most of his colleagues in the school kept extensive documentation on each of their students. It often seemed to Phil that these teachers spent the better part of their time correcting papers, entering marks into databases, and averaging long columns of numbers. Even after all that work, these teachers still wound up with a single grade that was supposed to represent a student’s work for an entire grading period.

It didn't seem right to Phil that two or three months' effort could be boiled down into a single grade. How was a
student to know what things needed improvement or what areas of his or her work were going well? Wasn't a teacher supposed to gain a fairly deep understanding of a student's educational development over the course of twenty to forty weeks, or several years? How could this sort of understanding be summed up in a B+ or a 78.5?

A Packed Schedule
Like that of many teachers in performance classes, Phil's schedule revolved around concerts, rehearsals, and extracurricular events, such as football games, parades, musicals, community festivals, and competitions. Although Phil enjoyed performing and believed that students, parents, and administrators all expected him to place a heavy emphasis on performance, he often grew weary of the concert/competition treadmill. He knew there was much to be gained from performing, but wondered whether he might be limiting the breadth of his students' experiences by focusing his efforts solely on playing ability.

Phil had been taught that music education included more than just learning how to sing or play an instrument. Listening, critiquing, analyzing, composing, improvising, creating—a musically educated person should be able to do all these things. "But there just isn't time," Phil thought to himself. "Even if I did all that stuff, how would I be able to come up with a grade that could accurately measure skills as different as improvising and analyzing or listening and composing? I'm having a tough time now, and all we're doing is rehearsing and performing."

The dilemma in which Phil finds himself is similar to the situation facing thousands of music educators every day. Already confronted with overloaded schedules, inadequate budgets, and a lack of support, music teachers are also being asked to redefine and justify their programs and sometimes to defend their jobs. In an effort to address these challenges, teachers, administrators, and researchers are joining forces to examine alternate methods for assessing both individual student achievement and overall program effectiveness.

Some New Terminology
The process of finding alternative assessment methods has added a number of new terms to the already cluttered educational lexicon. Here is a brief look at a few of these terms:

Performance-based assessment is a new name for something music teachers have been doing for years. Whenever a student plays an audition, sings a solo at a festival, or participates in an ensemble at a judged competition, he or she is engaged in a form of performance-based assessment. In this type of assessment, student performance is usually measured using some kind of rating scale. Teachers also use performance-based assessment techniques in evaluating student compositions, coaching fledgling jazz improvisers, and critiquing student-written reviews of concerts and recitals.

Traditional rating scales usually consist of a number of categories (such as tone, technique, diction, intonation) in which a student is assigned a number or letter rating based on a scale such as 1-2-3-4-5 or A-B-C-D-F. The disadvantage of this type of rating scale is that the distance between units is often unequal, causing those who make the judgment to arrive at their decisions based on subjective opinions of "goodness" for a given category. When, for example, at a festival, several adjudicators are asked to rate the same performance, the frequent lack of agreement in judges' scores can cause confusion and difficulty in interpreting the level of student performance and in determining which of the scores is the most accurate measure of an ensemble's or individual's true level of performance ability. Researchers who have studied the traditional music festival and have examined contest evaluations forms have found low levels of interjudge reliability, and the fairness and validity of such experiences has been questioned.2

Criteria-specific rating scales, on the other hand, are tools that are intended to help educators come up with somewhat more objective evaluations of performance-based activities. Two types of criteria-specific scales that are becoming more common are continuous rating scales and additive rating scales. Continuous rating scales list a series of increasingly difficult technical or musical criteria. Attainment of a particular rating is contingent upon
Figure 1. Continuous rating scale

Circle the highest level of achievement attained.

Student performance of the given four-measure excerpt:

5—was accurate throughout and performed with a consistent tempo.
4—contained one or two rhythmic errors and was performed with a consistent tempo.
3—contained three to four rhythmic errors and was performed with a consistent tempo.
2—contained five or more rhythmic errors and was performed with an inconsistent tempo.
1—was inaccurate throughout; performance lacked a consistent tempo.

successful accomplishment of all previous criteria (see figure 1).

Additive rating scales differ from continuous rating scales in that the “criteria descriptors”—the parts of the scale that describe the various elements of the performance being assessed—are not organized in a sequential manner, and attainment of a particular rating does not depend on successful achievement of all previous levels. Figure 2 shows an example of an additive rating scale.

Music teachers have long used rating scales similar to those shown in figures 1 and 2. Using precise language when describing assessment criteria can help the music teacher make the evaluation process more objective.

Interviews and Journals

Ensemble directors often find themselves dealing with large numbers of students, making the personal contact needed for in-depth learning to take place extremely difficult, if not impossible. Two techniques that may provide assistance in achieving better interpersonal communication in ensembles are interviews and journal-keeping.

Interviews can furnish a rich supplement to the teacher’s knowledge of individual students, enabling the teacher to reach a better understanding of students’ learning styles, interests, strengths, and weaknesses. This information can be valuable for educators who are trying to tailor their teaching to students’ individual differences. From Lyle Davidson and Larry Scripp, both of whom worked on Harvard’s Project Zero and are on the faculty of the New England Conservatory of Music, comes this list of additional advantages that the use of interviews can offer the educator:

1. Interviews make it possible to build lesson plans and curriculum on the basis of student’s actual knowledge as well as specific knowledge of the domain.
2. Interviews help teachers monitor the extent to which students integrate class work into their already active musical lives.
3. Through interviews, teachers can better assess the range of a student’s application of lessons and the degree to which students take or show initiative.
4. Interviews provide a ready and simple tool teachers can use for revealing the extent to which students in class or ensemble carry what they’ve been taught into playing, listening, and thinking about music.

Sometimes interviews can provide information about a student’s musical background that is unknown to the teacher. Here is a short excerpt from an interview with a student originally thought to be sullen, withdrawn, and noncommunicative about his musical interests: “Me and Theo make up songs on drums. ... We take turns picking the ones we like, play and practice them.... Sometimes we write separate songs at each house, then we play in my basement where my dad built a platform.” This snippet of dialogue reveals an active, vibrant musical life outside school that the teacher can use to establish contact with the student in the classroom.

To make interviews an effective part of the assessment process, teachers need to learn techniques for documenting and dissecting these dialogues. Good interviews can be time-consuming; the greater the teacher’s facility in conducting the interview, the more useful the experience will be for all concerned.

Journals give teachers another means of obtaining information about their students, especially information that might not come to light in a performance-based class like band, orchestra, or chorus. Frank Abrahams, of Westminster Choir College in Princeton, New Jersey, suggests using stenographer notebooks for student journal-keeping. These notebooks have the advantages of being easy to store and having pages that are separated into two equal, vertical halves, perfect for recording student thoughts on one side and teacher feedback and responses on the other.

An idea borrowed from the whole language movement, student journals offer a powerful vehicle for improving student-teacher communication, providing insights into individual attitudes, and assessing students’ understanding of curricular goals. While the initial forays into journal-writing can be quite structured (“Write two to three pages on someone who has been an influential person in your life as a musician’’), the real strength of the journal comes from its flexibility and open format. The blank pages of the journal await the students’ questions, hopes, and dreams. After students have been working on their journals for some time, the teacher might ask students to suggest three ways in which individual or ensemble performance could be improved and to suggest a number of techniques that could be used to effect these improvements. Added benefits of an activity such as this “ideas for better learning” assignment are the increased sense of
ownership students start to feel for their role in the learning process and the improvement in ensemble members’ attitudes, behaviors, and performance that can result from this sense of ownership.

Portfolios
From the visual arts comes another recent trend in evaluation, portfolio assessment. Various forms of this approach to measuring and evaluating student learning outcomes are currently being tried and tested in schools in such diverse regions as England and the Northeastern United States.6 But the application of portfolio assessment to arts education evaluation that may prove most exciting and useful to music teachers is exemplified by the Arts Propel project, a collaboration of the Pittsburgh (Pennsylvania) Public Schools, the Educational Testing Service, and Harvard University’s Project Zero.

Pittsburgh’s Arts Propel project is built around the three ways that students learn about music: production, perception, and reflection. According to Harvard University’s Howard Gardner, codirector of Project Zero and author of Frames of Mind: A Theory of Multiple Intelligences, “We think artistic learning should grow from kids doing things: not just imitating, but actually drawing, dancing, performing, singing on their own.”7

Thinking about learning in this way is not new to most music teachers. Rehearsals, recitals, concerts—music teachers are experts in these kinds of learning settings and could function as leaders for the rest of the educational community when it comes to teaching hands-on learning. The problem occurs when the issue of assessment arises. As a profession, we don’t always seem comfortable measuring the outcomes of our teaching, and here is where portfolios may help.

Ross Miller, of Nazareth College in Rochester, New York, identifies four types of portfolios.8 Each type might be thought of as a different kind of picture or group of pictures. Each looks at student performance from a slightly different angle, and each provides a new and unique perspective on student learning.

The presentation/product portfolio is analogous to the traditional artist’s portfolio. It can be used by the student when he or she is seeking admission to a school or when interviewing for a job; it represents the individual’s best finished work. It is easy to see how this type of portfolio is useful in formulating a summative or final grade.

Performance-based assessment is a new name for something music teachers have been doing for years.

The product/performance portfolio is the “class picture” of the portfolio family. To assemble it, the teacher collects the same product (for example, a written composition or an audiotape of a playing exam) from all students at the same time, allowing the teacher to make judgments regarding overall student progress.

The formal group portrait is the program portfolio. It “shows the best work of a group of students from a particular program … [and] might be assembled by a music teacher to help in the justification of funding for a music program or [when] presenting a request for additional resources,” according to Miller.9 This kind of collection, while similar to the product/performance portfolio in that it includes samples of work from many students, differs in its intent. The program portfolio is meant to serve as a representation of exemplars of student work over time; the product/performance portfolio serves as a ranking and sorting instrument for class measurement and evaluation.

The final example, the process portfolio, may be thought of as a slide show, photo album, or scrapbook. It differs from the program portfolio in that it includes early “and (perhaps) less successful attempts at production.”10 The element of reflection, so crucial to the Arts Propel program, is of primary importance in the process portfolio. Students are encouraged not only to create, but also to revise, reevaluate, and refine. The creative process is seen clearly through a student’s halting efforts as a producer of music, in much the same way that one

---

**Figure 2. Additive rating scale**

Check all those that apply.

The student demonstrated:

- characteristic tone quality
- appropriate dynamics
- appropriate articulations
- consistent tempo
- acceptable intonation
- Add number of checks; enter here.
can observe the evolution of a piece of music by examining a professional composer’s sketchbooks.

Technology can play an important role in the development of process portfolios. The revision process that is critical to composing music can be accomplished easily using the data-recording features that are part of much music sequencing and printing software. Students can hear their compositions immediately; they can reassess their work and make revisions with relative ease through the computer’s MIDI capabilities.

Process portfolios also “enable students to take ownership of their work, typically low achievers as well as more gifted students.” For most teachers, it is this feature of alternative assessment that seems to hold the greatest promise. Educators often wonder how best to evaluate students who are in the same class yet possess vastly different abilities, intelligences, learning styles, backgrounds, motivations, and attitudes. As Phil, the frustrated music teacher, has thought many times, “There must be more than a hundred points between a senior auditioning for music school and a freshman taking band just to get the one unit of credit in the arts needed for graduation! How can I come up with grades that are meaningful for both students?”

Back to Square One

The first step that Phil needs to take is deciding what kinds of musical behaviors he wants from his students and how best to measure and evaluate their understanding of the concepts associated with these behaviors. These decisions are different for each teacher and depend on a variety of factors, such as facilities, budget, schedule, administrative and parental support of the music program, and curricular philosophy. What is common to all situations, however, is the symbiotic relationship between the teaching and evaluating components of the learning process. Alternative methods of assessment can, according to Frank Abraham, “promote authentic learning: higher-order thinking, depth of knowledge, connection to the world beyond the classroom, substantive conversation, and social support for student achievement … [and] document authentic musical behaviors including, but not limited to, performing, practicing, reflective thinking, listening, composing, improvising, studying, and researching.”

Whenever a student plays an audition, sings a solo at a festival, or participates in an ensemble at a judged competition, he or she is engaged in a form of performance-based assessment.

Once Phil has decided to give alternative methods of assessment a try in his classroom, he needs to think about what types of materials should go into a portfolio. This can lead to the most daunting part of the evaluation procedure: where to start. A few key thoughts to keep in mind, based on Edwin Gordon’s work, are the following:

- You should be able to measure what you expect your students to learn.
- Students should be asked to function at multiple levels of learning, including both discrimination and inference levels. (Discrimination learning is rote learning, and it involves perception, sensation, and audition. Inference learning involves making generalizations from familiar to unfamiliar information. Discrimination learning provides the necessary readiness for inference learning.)

Using these two prerequisites as gatekeeper considerations can prevent the portfolio from becoming merely a scrapbook of children’s musical refrigerator art and can help it become a powerful tool for collecting student efforts at product and process learning in music.

Some of the types of materials to be found in portfolios, therefore, may include “student journals, rehearsal tapes, aptitude and achievement test scores, playing or singing examinations, listening assignments, ensemble rehearsal critiques, error detection musical scores, domain projects [borrowed from the Arts Propel project], self-evaluations, contest scores, audio- and videotapes, concert reviews, graded items, ungraded items, [and] mid-year and final examinations.” This list includes examples of both authentic assessments (student journals, audiotapes) and traditional assessments (standardized tests, aptitude tests, and achievement tests). This inclusive style of assessment is referred to by Michael Kean, vice president of Macmillan/McGraw-Hill and chair of the Test Committee of the Association of American Publishers, as the “multiple-measures approach,” and it seems to answer some of Phil’s questions with respect to grading the wide variety of students he finds in his band each year.

For example, Phil began to see how an audiotape of students playing exams, added to at regular intervals throughout the school year (an example of an ongoing product/performance portfolio activity), could make his grading process much easier and more valid. He devised a continuous rating scale to evaluate each student’s performance, and he listened to the tapes in batches throughout the marking period in order to arrive at scores. He thought it might even be worthwhile to have each student do a self-evaluation of his or her own tape and use both his score and the student’s score as part of the performance grade. “Maybe,” Phil thought, “they will take more ownership of the grading process if they have a meaningful voice in the results.”

Making Real Music

Interested in including measures that would help provide a fuller, richer picture of student learning, Phil began to think about designing a project that
would allow his kids to experience what real musicians do—perform, compose, and listen to music—using the same types of assessment processes that professional musicians use in their daily activities. Phil knew that much of the action in his band room was very teacher-directed and thought that his students might show more interest and motivation if they felt more involved in the planning of the teaching/learning process. He also thought that having kids work together in small groups was something he had not done enough of and something that could help his students become more a part of the classroom learning experience.

The project that Phil eventually put together involved much more than a typical performance. It required his students to organize a chamber ensemble (such as a brass quintet, a flute trio, or a percussion ensemble), select an appropriate piece of music, study the historical background of the composer and the period in which the piece was composed, and prepare the music for performance. Phil volunteered to help each ensemble find an audience for which to perform (an elementary music class, a high school social studies class studying the historical period in which the piece was composed, or a senior citizens' club were all examples of groups that had requested this type of performance from Phil's students in the past). He offered to help the various ensembles prepare a program that would include not only playing the piece but also a short presentation of the information they had uncovered in their investigation of the composer and the historical period. This combination of performing and informing is sometimes called an “informance.”

Members of the ensemble were encouraged to participate in the informance in a manner in which they felt comfortable. Students at ease in front of a group were assigned (by the ensemble members) as speakers; those who were more reticent took responsibility for preparing handouts or overhead transparencies, writing program notes, or handling other research duties before the performance. To make the evaluation of the project more meaningful for the students, Phil asked them to help create the standards by which to measure their own musical growth and performance. He was surprised at how much tougher the students were on themselves during the project than he would have been, and found himself reminding the groups how much they had learned. He seemed to be spending less time criticizing their performances. Phil was starting to feel more enthusiastic about what was happening in his band room, and he even began to feel more confident that the next marking period's grades would be meaningful for both himself and his students.

Criteria-specific rating scales ... are tools that are intended to help educators come up with somewhat more objective evaluations of performance-based activities.

A Leap into Composition

So far, the student portfolios included audiotapes of student playing exams and evidence of their chamber music projects, both of which were performance-based assessments. Phil also wanted his students to have experience with some of the other forms of musical learning, such as composition or analysis. At the same time, Phil was unhappy with the quality of much of the music currently available for his ensembles, and he wondered whether there might be any budding composers in his band. He reasoned that his students might be more interested in practicing music they had written themselves than music that had been purchased from a store, and so he began to develop a student composition program.

Phil started the first class by surveying the students about what kinds of pieces they wanted to write, and found that most of them were interested in writing music that they could play with their friends. After introductory lessons on orchestration, transpositions, ranges and tessituras, and instrumental combinations, the students broke up into small groups of three to eight students. Phil roamed from group to group, offering advice and suggestions as requested, but spent most of his time watching and listening.

As the student compositions neared completion, Phil began to prepare the students to teach their pieces to the groups for which they had been written. Conducting patterns and gestures were demonstrated, and rehearsal techniques discussed. The student performers were asked to critique both the new works and the manner in which they were presented by their composers, a valuable experience shared by too few musicians at any level. While some of the compositions needed substantial reworking, most were surprisingly well-crafted, and these works supplied Phil with an expanded library of chamber music and a roomful of satisfied young composers.

An Improved System

With the marking period looming, Phil scheduled appointments with his students to look over the contents of their portfolios. With so much tangible evidence of their work, the students had a pretty good idea of what grade they would be receiving, and there were no surprises when the quarter's report cards were issued. Not every student was happy with his or her grade, but there were fewer complaints and more positive attitudes in rehearsals and lessons than there were at the close of the previous marking period. Phil felt good about the learning that had taken place in his band over the course of the quarter, and had already come up with more ideas for
projects for the following year.

Phil realized, however, that there were a few problems regarding these new forms of assessment he still needed to address. First, the new approach took a lot of time! Phil was taking home journals several times a week, listening to student tapes on his lunch break, and grading tall stacks of assignments almost every weekend. Added to the evening and weekend commitments he already had scheduled, Phil found himself with no time to think, much less to reflect and revise. He was also running out of room. His tiny office was now home to several large boxes, crammed with expandable file folders brimming with tapes, journals, and scores. The practice room/repair shop/storage closet that he shared with the orchestra and chorus teachers was so overstuffed that he was afraid to open the door.

The question of how to convert these stacks of information into a term evaluation or final grade for each individual student was still unanswered. Phil knew that he could convert each project, composition, or audiotaped playing assignment into a numerical mark through the use of criteria-specific rating scales, but thought that he might be missing the whole point of alternative assessment in doing so. Unfortunately, the principal would still be expecting Phil’s grade sheets (with all the little bubbles neatly penciled in) to be delivered to the main office with those from all the other teachers. Phil took some comfort from the feeling that he was beginning to know more about his students than ever before, and he believed that this knowledge was making him a better teacher.

Phil was still thinking about these problems when he ran into his former graduate-school colleague, Cathy, at a county music teachers’ meeting. After they exchanged pleasurabilities, Phil discovered that his colleague had been experimenting with portfolio assessment for the past year and had run into similar time and space difficulties. Cathy suggested that Phil might have bitten off a bigger chunk of the assessment pie than he could chew. She advised Phil to have his students work on only one or two projects at once and to stagger assignment due dates across grade levels to reduce the avalanche of paper in which he found himself buried. She also suggested that having the students store their written assignments on floppy disks might further reduce Phil’s paper flood. Just knowing that someone else had gone through similar growing pains made Phil feel a little more hopeful.

You should be able to measure what you expect your students to learn.

A Fresh Look

Phil’s experience is a good example of both the pros and the cons of alternative assessment. While these new ways of measuring student learning can provide teachers with fresh perspectives on their own teaching styles, the commitment of time and energy necessary to allow these methods to work effectively can prove demanding. Moderation in all things is the key. Phil may find that Cathy’s suggestions to limit the number of projects and stagger student due dates will make his workload more manageable. Giving students more responsibility for their own evaluations may also help Phil handle his new duties more efficiently.

For many teachers, alternative assessment is just a new name for what they have been doing all along: using their musical and pedagogical talents to help their students experience the joys that music has to offer in an efficient, yet enticing, fashion. When portfolios are used without thoughtful planning and continuous revision, they become merely fancy scrapbooks, no more or less valid or reliable than any other form of measurement or evaluation can be. In the final analysis, it is not what is in the portfolio that matters; rather, it is what is inside the teacher and the student.

Notes


4. Davidson and Scripp, “Arts Propell Portfolio Interviews,”


