How can student learning in the arts be assessed?

Chelsea Kari
March 10, 2014
Introduction

A consistent theme across arts education literature is the difficulty of measuring student learning in the arts. Assessing arts learning is important for a number of reasons, including measuring the effectiveness of arts education programs and demonstrating the importance of the arts in school curriculum. In spite of its importance, assessment in the arts is frequently disjointed and random. The arts education community has struggled to agree upon how arts learning should be measured, and even what should be taught. Although a standardized test for the arts was attempted, and there have been *National Standards for Arts Education* since 1994, assessment and program content is mostly left up to individual arts teachers. Thus, there are a number of assessment strategies being used. Teachers employ performance assessments, portfolio assessments, and self-evaluation by students to maximize learning outcomes and measure student learning. There is no one correct method for assessing student learning in the arts.

Why Assessment is Important

Assessment of student learning in the arts serves a number of important functions. From the perspective of national, state, and local governments, school districts, and arts education funders, assessment holds instructors accountable for teaching to the standards. Assessment also provides a tool for comparing students to one another (Castiglione, 1996, p. 4). Accountability-based testing is a great source of contention in the arts education community.

Another important function of assessment is shedding light on how arts education programs can be improved. Measuring how much students have learned provides insight into how well the material is being taught. Information gained from assessment should
“indicate where arts education needs to be strengthened and extended so future students can fully realize their potential” (National Assessment Governing Board, 2008, p. 3).

Assessment can have a positive impact on student learning. Rabkin, et. al (2011) write that good assessment “advances students’ inclination to assess their own learning, and challenge themselves” (p. 108). When assessment of student learning is ongoing and occurs throughout the learning process, it has the ability to strengthen final performances or art works (Wolf & Pistone, 1995, p. 46). Assessment also allows teachers inheriting students to know where those students are in their learning journey (Castiglione, 1996, p. 3).

One of the most important roles of assessment is supporting arts education advocacy. Assessment provides a tangible representation of the positive outcomes of arts learning to parents, school district boards, and policymakers. According to the National Assessment Governing Board (2008), “Assessment has the unique ability to fix attention in education” (p. 2). Subjects which are tested are perceived as more legitimate because there are measurable indicators of student learning. Arts assessment can help provide evidence that the arts should play a larger role in K-12 public schools.

The Problem

The arts education community has struggled to agree upon proper means of assessing student learning. Because of this, “arts assessment in most of the United States consists of individual teachers developing and using their own assessments to suit their own needs” (Bodilly, et. al, 2008, p. 15). Learning in the arts is more difficult to measure than learning in other subjects. As the National Assessment Governing Board (2008) points out, assessment in the arts “is complex because of the difficulty in constructing an
assessment that accurately appraises student achievement on a national level with all the variables of experience and environment, and delivers it in a timely, cost-efficient manner” (p. 6). Finding a way to measure something as intangible as arts learning is one of the main challenges facing the arts education community.

In *Arts Education: Defining, Developing, and Implementing a Successful Program* (2003) Senior Director of Arts Policy at Americans for the Arts Marete Wester describes the inherent difficulty in “measuring something that resists quantification” (p. 182). Art is unlike other many other academic subjects, in that there is no one right answer or correct way of doing things (Wolf & Pistone, 1995, p. 8). People’s opinions of works of art are necessarily subjective; two individuals can have completely different reactions to the same performance or artwork. Much of the arts experience is a matter of taste and personal preference. A significant concern in the arts education community is that “an assessment of the arts will artificially quantify those essential aspects of the art that seem unquantifiable” (National Assessment Governing Board, 2008, p. 4). Creativity, which is integral to the art-making process, is not something that can be measured on a scale of one to ten.

Learning in and through the arts has been associated with higher levels of achievement in other academic subjects, the ability to think critically and problem solve, greater self-confidence, and the development of social skills (Ruppert, 2006, p. 10). Assessment attempts to measure those positive outcomes, therefore demonstrating the importance of arts learning to policymakers and school officials. However, “Establishing a causal link between any sort of educational intervention, like arts education, and positive student outcomes is profoundly tricky” (Rabkin, et. al, 2011b, p. 107). It is
nearly impossible to control for outside influences in order to prove that arts learning is in fact responsible for those benefits. Stanford Emeritus Professor Milbrey McLaughlin (n.d.) writes that, “Meaningful [evaluation] measures acknowledge that many outcomes important for youth to achieve—confidence, agency, leadership, responsibility—are difficult to assess, especially in the short run” (p. 24). Many of the benefits of arts learning manifest themselves over time, and therefore cannot necessarily be measured through in-school assessments.

There is disagreement as to what assessment tools can accurately measure arts learning. In Qualitative Assessment of Arts Education (2008), Robert Stake and April Munson write that, “Conceptualizing quality in arts education with scales and rubrics can be problematic. Student performance is more complex than any checklist that a teacher or assessment expert can make” (Stake & Munson, 2008, p. 16). Standard assessment tools such as multiple-choice paper-and-pencil tests are not going to be effective in measuring student learning in the arts. Even more qualitative measurements of learning, such as performance assessment, cannot capture all the nuances of student achievement in the arts.

Ways of Measuring Learning

There is a great deal of variety in the methods arts educators and assessment professionals use to measure arts learning. Achievement in arts learning is frequently measured in terms of student behaviors, like attendance and participation. Other arts educators argue that, “Careful analysis of student work is the best way to understand what they are learning” (Rabkin, et. al, 2011, p. 101). Analyzing student work is not cut and dry, however. Within the analysis of student work, there is disagreement on what
tools should be used, when and how often the assessment should be conducted, and by whom. In the following section, a few of the multitude of assessment techniques are illuminated.

**Quantitative vs. Qualitative Assessment**

There are both quantitative and qualitative forms of assessment. In the arts, “Quantitative assessment is scalar, multidimensional, and based on criteria” and “usually emphasizes expressed standards, comparability, objectivity, and the social science of aesthetics” (Stake & Munson, 2008, p. 14). Because it is based on clearly defined predetermined criteria, quantitative assessment is viewed by some to be more reliable and consistent than qualitative assessment. According to Stake and Munson (2008), “Qualitative assessment is fundamentally personal, rooted in the comprehensive experience of the person or persons evaluating” (p. 14). The personal nature of qualitative assessment is sometimes seen as a weakness, allowing assessment to be tainted by personal biases. On the other hand, “Qualitative assessment is attractive to many because it is holistic, contextual, empirical, and empathic” (Stake & Munson, 2008, p. 14).

Because the arts are nearly impossible to quantify, quantitative assessment cannot capture many of their more ephemeral aspects. Allowing for subjectivity gives assessors the ability to account for those hard-to-measure traits. Therefore, the best approach to assessment likely combines quantitative and qualitative assessment characteristics. In this type of assessment, “Performance is rated judgmentally but with a scaled rubric, alluding to quality, often in a disciplined way, and with comparisons” (Stake & Munson, 2008, p. 16).
Taking Full Measure authors Dennie Palmer Wolf and Nancy Pistone (1995) outline the following characteristics of good assessment: an insistence on excellence, judgment, the importance of self-assessment, the use of multiple forms of assessment, and ongoing assessment (p. 8). Arts learning encourages students to take risks and constantly strive for higher levels of achievement. Assessment of arts learning needs to reflect that. Art-making consists of a series of decisions by the artist. Assessment of learning in the arts should assess all those individual decisions, not just the end result. Student involvement in assessment is the best way to ensure that students from assessment instead of just receiving a grade or a test score. The use of multiple forms of assessment captures more aspects of arts learning than any one assessment. Finally, ongoing assessment provides students with an opportunity to take feedback into account and use it to inform their future work.

**Indicators**

In Program Evaluation: Measuring Results Craig Dreeszen classifies types of assessment according to what sorts of indicators they use. Although Dreeszen is specifically focused on program assessment, the types of indicators he describes are useful in assessing student learning as well. Dreeszen (2003) defines indicators as “measurable evidence that shows the extent to which an intended outcome has been achieved” (p. 255).

The first type of indicator Dreeszen identifies is physical evidence, or student artwork (2003, p. 271). It stands to reason that student learning should be measured by the work students produce, whether that takes the form of paintings, poems, recitals, or performances.
Another type of indicator is participant answers to questions collected through “interview, focus groups, evaluation forms, questionnaires, [or] tests” (Dreeszen, 2003, p. 271). These indicators are useful in providing evidence of program success. However, they can also be used for measuring how well students are understanding material. Having students actively evaluate their own learning is just as important as evaluation by an outside party.

Evidence of student learning can be collected through “observation by a skilled teacher, evaluator, program coordinator, or participant observer” (Dreeszen, 2003, p. 271). When done well, observation of student learning is conducted in a way that ensures consistency and reliability. This entails using a predetermined rubric and recording observations at regular intervals throughout the arts education program.

The final indicator Dreeszen identifies is archival information, which includes “attendance records, reported changes in school disciplinary actions, [and] test scores” (Dreeszen, 2003, p. 271). This type of indicator serves as a measure of accountability to school officials and funders. Such information is frequently used in arts education advocacy. However, attendance records and test scores do not provide a full picture of the impact of arts learning.

**Student Behavior**

Student learning can be evaluated based on its impact on student behaviors. In their study of Teaching Artists (TAs), Rabkin, et. al (2011) noted that TAs “evaluate engagement by whether students contribute to discussions, enjoy and apply themselves, reflect on and try to improve their work, and listen and show interest in each other’s work” (p. 10). Behaviors such as these are difficult to capture in a test. These behaviors,
however, are more likely to contribute to student success than their ability to answer a series of multiple-choice questions correctly. Teaching Artists were able to tell that their students were engaged “through visible and physical indicators – from their postures and the focus in their eyes, to their active participation in warm-ups and games, to the energy they invest in doing their work” (Rabkin, et. al, 2011, p. 101). Such observations are more meaningful indicators of student growth than grades and test scores.

**Linguistic Assessment**

One unique way of measuring the impact of arts learning is the analysis of linguistic patterns among students. In a decade-long study of 124 youth-based community organizations, Milbrey McLaughlin and linguistic anthropologist Shirley Brice Heath found a link between participation in youth-based arts organizations and the level of complexity of language used by students. In *Champions of Change* (n.d.), Heath and Roach noted, “The influences of participation in the arts on language show up in the dramatic increase in syntactic complexity, hypothetical reasoning, and questioning approaches taken up by young people within four-to-six-weeks of their entry into the arts organization” (p. 27) Linguistic patterns observed included increases in the use of if-then statements, mental state verbs, and modal verbs (Heath & Roach, n.d., p. 27). Heath and Roach make the argument that these patterns signal “cognitive and social growth” (Rabkin, et. al, 2011, p. 98). Tracking the sorts of language used by arts program participants is just one example of how to measure student learning. The authors of *A Report on the Teaching Artist Research Project* further noted that, “the language that Heath found prevalent in the arts programs is the language of artistic assessment: language that questions and tests, proposes and considers, accepts and rejects” (Rabkin,
et. al, 2011, p. 98). This provides an example of how integral assessment is to the artistic process.

**Performance Assessment**

Performance assessment is a type of assessment where students are required to carry out a specific task instead of regurgitating memorized information. It is sometimes called authentic assessment because it more closely resembles how work is assessed in the “real-world”. Similar to work evaluations, performance assessment in schools includes strategies like performances, critiques, and personal reflection (Saraniero, n.d.b, para. 3). Applied to the arts, “Performance assessment requires student to perform a task that results in a product (such as a sculpture or a composition) or a performance (such as a concert or a dance recital)” (Saraniero, n.d.b, para. 4). Students are judged both on the process of creating the work and the work itself. Advocates of performance assessment prefer it to paper-and-pencil testing because it requires students to “integrate their newly acquired knowledge and skills by doing the work (Saraniero, n.d.b, para. 5). Performance assessment is believed by many to provide a more complete picture of student learning than standard testing methods.

In *Portfolio assessment in art and art education* (1996), Lawrence V. Castiglione identifies several characteristics of performance-based assessment. Performance-based assessment uses open-ended tasks like writing a poem or performing a skit (Castiglione, 1996, p. 1). It “focuses on higher order or complex skills” as opposed to route memorization (Castiglione, 1996, p. 1). Performance assessment “employs context sensitive strategies” by attempting to recreate an authentic art-making environment (Castiglione, 1996, p. 1). It “often uses complex problems requiring several types of
performance and significant student time” (Castiglione, 1996, p. 1). By taking the time to observe students in an environment typical of the arts discipline, assessors are able to get a better idea of what a student knows. Finally, performance-based assessment “consists of either individual or group performance” and “may involve a significant degree of student choice” (Castiglione, 1996, p. 1). The aforementioned factors combine to make performance-based assessment a more natural, less stress-inducing strategy for measuring student learning.

Performance assessment is tailor made for use in the arts. As noted by the National Assessment Governing Board (2008), “Performing is impossible to assess through any means other than the processes themselves” (p. 19). A student sitting down and taking a multiple choice test does not capture the full range of learning in the arts. University of San Diego Instructor Patti Saraniero (n.d.b) argues that, “Engaged and active instruction is best assessed by engaged and active assessment.” (para. 7). As a result, performance assessment is now used in many non-arts disciplines because of the comprehensive picture of student learning it provides. In 1996, Lawrence V. Castiglione observed that, “Performance-based assessment methods are now used in education more widely than at any time since the multiple-choice test format was introduced in about 1914” (p. 1). The use of performance-based assessment in schools is frequently advocated for by education reformers (Castiglione, 1996, p. 1).

Performance assessment is not embraced by all. Critics argue that performance-based assessments are “much less efficient, much more time consuming, less easily standardized, and obtain smaller samples of pupil performance” (Castiglione, 1996, p. 7). Compared to tests with only one right answer, performance-based assessment is
subjective in nature. Performance assessment leaves room for the biases of evaluators to influence results. Performance assessment is also, as Castiglione notes, incredibly time consuming because it involves observing the process of making a final product instead of just judging the final product itself. Just like all assessment methods, performance assessment is imperfect.

**Portfolio Assessment**

A common form of performance assessment is portfolio assessment. The National Assessment Governing Board (2008) defines a portfolio as “a collection of work produced over time and unified by a theme or purpose” (p. 28). Portfolios can be either summative or formative. A summative portfolio includes only finished pieces of work, and usually includes what a student perceives to be their best pieces of work. Formative portfolios include drafts and works in progress, providing greater insight into students’ mental processes. Formative portfolios “provide illustrations of the difficulties and problems students encountered while they were learning” (Castiglione, 1996, p. 4). Because of this, “Portfolio assessment has potential for providing rich source of information that cannot be gleaned from most conventional assessment techniques” (Castiglione, 1996, p. 3). Advocates of portfolio assessment would argue that capturing the student’s thinking process is more important than measuring how many answers a student got right or wrong.

While there are standardized ways of evaluating student portfolios, portfolio assessment is subjective in nature. Usually a teacher or a panel of judges will use a rubric or checklist to assess student portfolios (Saraniero, n.d.a). An example of formal use of portfolio assessment is the Advanced Placement Studio Art Portfolio Evaluation, which
“asks students to submit collections of original works that are assessed holistically by groups of raters” (National Assessment Governing Board, 2008, p. 28). It is important that the assessment of portfolios is as unbiased and objective as possible.

Standardized criteria and training of judges can help reduce bias in portfolio assessment. However, as Castiglione (1996) notes, “Training portfolio assessors to achieve consistency, fairness, and accuracy requires time, effort, and expensive field trials” (p. 6). Because of this, and legitimate concerns about accuracy, conducting portfolio assessment on a large scale is usually not attempted.

**Ongoing Assessment**

Ongoing assessment is favored by many arts education experts, including Dennie Palmer Wolf and Nancy Pistone, who view assessment as an “episode of learning” (Wolf & Pistone, 1995, p. 9). As mentioned earlier, when assessment occurs throughout a class or program instead of at the end, students are able to internalize feedback and use it in their work. During ongoing assessment, students critique their own work and receive feedback from their teachers and other students. Rehearsals provide a great opportunity for ongoing assessment. In rehearsals, students get to try out different techniques and take risks. They get feedback on what they are doing and continue to mold their performance accordingly. Wolf & Pistone (1995), write that students should “see rehearsal as a place where they should be constantly noticing, making judgments, and raising questions” (p. 34). In this way, assessment becomes embedded in the learning process. In ongoing assessment, teachers evaluate students on “more than the quality evident in a final piece or performance” (Wolf & Pistone, 1995, p. 61).

**Standardized Testing of Arts learning**
In spite of widely held reservations about standardized testing, in 1992, the National Assessment Governing Board awarded a contract to the Council of Chief State School Officers (CCSSO), the College Board, and the Council of Basic Education to begin planning an arts test to be included in the National Assessment of Educational Progress (National Governing Board, 2008, p. iv). Bodilly, et. al (2008) argue that although this assessment rapidly fell into disuse, “it represents an important step forward in the development of valid and reliable tests for the arts” (p. 15). Those designing the assessment worked closely with those developing the 1994 National Standards for Arts Education. Both the assessment and the standards address learning in the areas of dance, music, theatre, and the visual arts.

The committees developing the assessment framework tried to balance the need for measurable standards and accountability with the nature of the artistic process. The assessment steering committee laid out guidelines including, “The assessment should assess students’ knowledge, attitudes, and performance in the modalities and forms of expression characteristic of the arts (music, dance, painting or drawing, acting) as well as verbal or written linguistic modes, that is, writing or talking about the arts” (National Assessment Governing Board, 2008, p. vii). The assessment attempts to measure three processes: Creating, Performing/Interpreting, and Responding (National Assessment Governing Board, 2008, p. 9) through “multiple, related exercises organized around an activity” (National Assessment Governing Board, 2008, p. 5).

The assessment was designed to be taken by students in the 4th, 8th, and 12th grades, and covers grade-level appropriate content in each of the four artistic disciplines. According to the National Assessment Governing Board (2008), “The assessment
exercises will consist mostly of constructed responses along with some multiple-choice items…Some constructed responses will ask students to perform using the language of the art form…Others will employ some short or extended written responses” (p. 23). Most of the assessments consist of a student observing a work of art, responding to and analyzing that work of art, and then creating a work of art of their own (National Assessment Governing Board, 2008, p. 23). Based on their performance, students would be judged as basic, proficient, or advanced (National Governing Board, 2008, p. 39) Those achievement levels “are framed by the three major arts education processes—creating, performing, and responding—and are set at the three grade levels examined” (National Assessment Governing Board, 2008, p. 39).

In planning the assessment, the committee ordained that, “the exercises must offer the same opportunities and the same challenges, and should be available in the same circumstances, for all the students assessed” (National Assessment Governing Board, 2008, pp. 23-24). Ensuring that students across the country take the test in similar and equal circumstances is extremely problematic. Factors such as available space and available equipment necessarily impact the test-taking experience. Another concern that arose concerning the assessment was the issue of fair judges. In anticipation of this problem, the planning committee declared that, “Facilitators who will conduct this assessment in the schools must be sensitive to the creative process and must possess some understanding of the arts discipline being assessed” (National Assessment Governing Board, 2008, p. 26). It would be inappropriate for individuals without a background in the arts to judge student performances. The significant challenges posed by such an assessment explain why it has only been used twice.
Why Standardized Testing is Problematic

There is significant disdain for standardized testing in the arts education community. Rabkin, et. al (2011) argue that,

Assessment in the arts defies the essential characteristics of standardized testing, the prevailing school-based assessments. The arts value subjectivity, originality, and the creation of integrated whole products, while the standardized tests so widely used in schools require objectively right answers and a focus on particular bits of knowledge. (p. 99)

Other arguments against standardized testing include that they encourage students to memorize facts instead of understanding big concepts (Wester, 2003, p. 182). Many worry that standardized testing can’t pick up on the many nuances of student performances and works of art. Art is too subjective to be measured quantifiably. Others argue against standardized testing because it provides no benefits to the students in the form of feedback on how to improve performance (Wolf & Pistone, 1995, p. 75). Stake and Munson (2008) summarize these concerns, writing, “Standardization may streamline the management of arts education, but it diminishes the role of the teacher. It is a step away from acknowledging the complexity of quality” (p. 16).

Conclusion

The topic of arts assessment is incredibly complicated and controversial. Art does not lend itself to being neatly measured and evaluated. Each individual will experience art differently, and what makes a good or a bad performance is frequently difficult to agree upon. Rubrics and checklists can provide some structure to arts assessment, but truly consistent, unbiased evaluation is hard to come by. In spite of these difficulties, the arts education community should continue to seek new ways of assessing arts learning. Assessment is too important to the cause of arts education advocacy to ignore. While a
nationally administered standardized test may not provide the answer, arts educators should continue to measure and document student learning in the arts.

References


