BY DONALD D. GRUBER

Measuring Student Learning *in Art Education*

I n this article, I explore the idea of measuring student learning in art education and endorse a method of assessment for use by art educators. My interest in assessment began early in my teaching career. I became aware then that there was a lack of emphasis among art educators toward assessment. For many of them, the idea of assigning grades to children's artworks presented an ideological dilemma. They viewed assessment as subordinate to the significance of children's creative experiences. But students *do* learn while creating and that learning *can* be measured.

Historical Overview of Assessment in Art Education

Until relatively recently the principal emphasis within the field of art education was on creative development with little regard to any substantial measurement of the learning that accompanied that development. However, as early as 1926, psychologist Florence Goodenough developed the *Draw a Man Test* in an effort to measure intelligence through drawing (Goodenough, 1926). The *McAdory Art Test* of 1929 and the *Meier-Seashore Art Judgment Test* of 1930 were two attempts to determine artistic ability (Gaitskell, 1958; Meier, 1966; Gaitskell & Hurwitz, 1970). Neither of those endeavors, however, accounted for actual learning in art.

Art education textbooks written prior to the 1960s made little or no mention of assessment procedures. Some art education writers advocated the advancement of expressionistic creativity. Others promoted the mental and physical growth potential of exposure to art processes. Assessment of student learning in art education was, for the most part, not a central focus. This earlier lack of emphasis on assessment in art education may have stemmed from the prevailing belief that art was not considered an academic discipline and, therefore, did not require the learning and assessment practices of core academic subjects.

Observation strategies involve consistent surveillance of what students actually do during all phases of an art lesson: performance, attitude, work habits, and behavior in general. (Students Andy Siharath, seated, and Erica Reese.
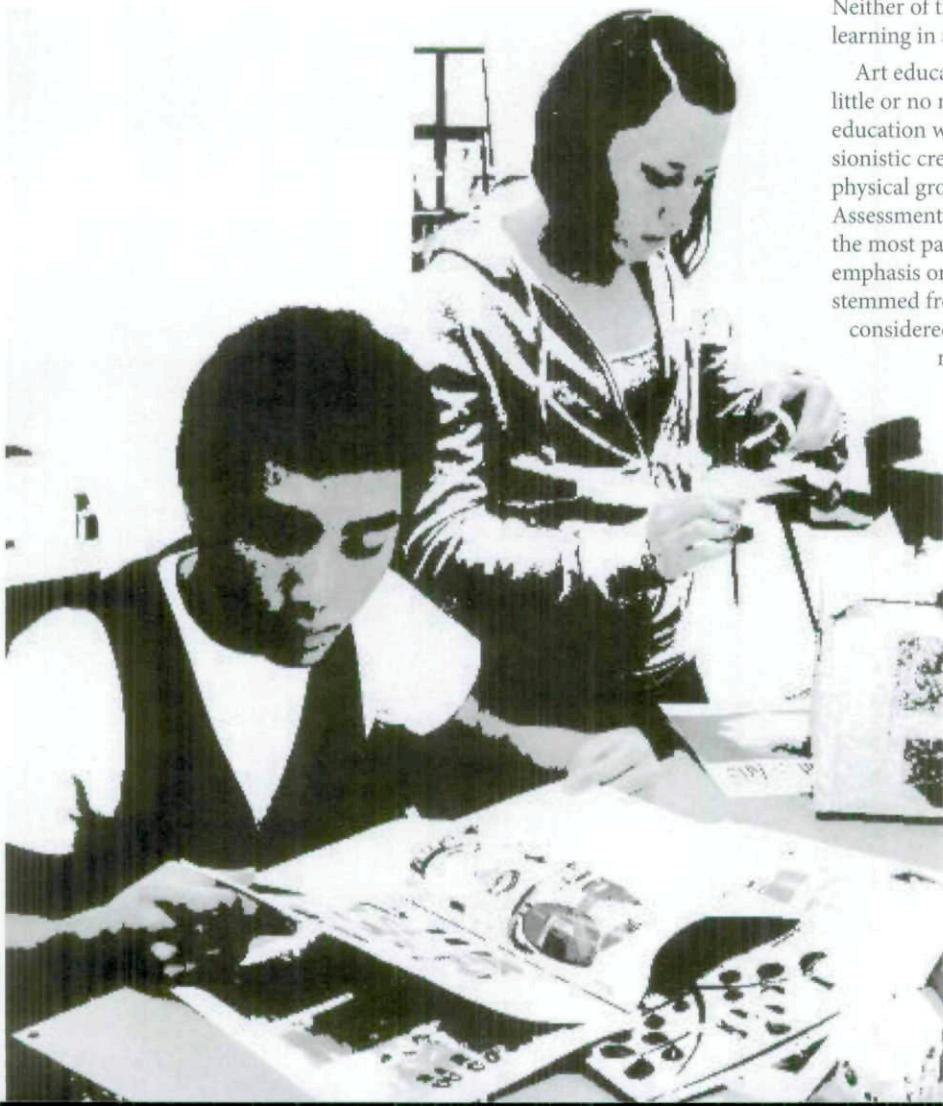
According to the National Art Education Visual Arts Standards, **evaluation** encompasses the global aspects of the curriculum. **Assessment**, on the other hand, refers to more tightly focused measurements at the level of the individual student and his or her interactions within the art program. Assessment, therefore, measures student learning and evaluation measures program efficacy.

By the late 1950s and early 1960s, American education was forced into a new direction, which emphasized rigor, quality, and measurement. Some attribute this change to the Russian space program and its early accomplishments. The real reasons, however, were far less simplistic and were already set in motion long before Russia's successful launch of Sputnik in 1957.

An interesting mix of philosophies regarding assessment of learning in art had already begun to appear in the literature. By the early 1960s, educators Manuel Barkan, Elliot Eisner, and Ralph Smith were advocating the idea of structure and discipline in art education (Barkan, 1962; Eisner, 1966; Smith, 1966). Moreover, by the mid-1980s, art education was firmly considered a distinct academic discipline (Efland, 1990).

However, in the art classroom, assessment still did not receive significant emphasis. Nonetheless, recent calls for accountability along with current trends toward academic standards have combined to create an interest on the part of art teachers in assessment strategies.

The Importance of Measurement of Learning

The classical approach to determining grades in art found the terms *assessment* and *evaluation* used interchangeably. However, the two terms have now been clearly defined as to their appropriate use and context. According to the National Art Education Visual Arts Standards, *evaluation* encompasses the global aspects of the curriculum. Assessment, on the other hand, refers to more tightly focused measurements at the level of the individual student and his or her interactions within the art program. *Assessment*, therefore, measures student learning and evaluation measures program efficacy. For the purposes of this article, assessment will be the primary focus of measurements of student learning.

Educational Accountability: It's the Law

Currently, educational accountability underscores the need for reliable assessment and evaluation to support innovations in curriculum design, instructional methods, program funding, and the appraisal of student achievement. In addition, there are pressures associated with the No Child Left Behind legislation.

Developing Assessment Criteria

One objection to long-established art assessment strategies rests with their subjective nature. Some teachers evaluate art products based on vague categories without regard to specified criteria that help determine academic growth and learning (Day, 1985). Those methodologies tend to be highly subjective and are generally based on personal aesthetic biases. Several writers have observed that this subjectivity can be reduced if explicit assessment criteria compatible with lesson objectives are developed at the outset of the lesson planning process (Gruber, 1994; Hardiman & Zurnich, 1981; Lowenfeld & Brittain, 1975; Rush, 1987, 1989). Objective outcomes must conform to specifically described criteria developed expressly for a given lesson.

Two types of assessment criteria are described in the literature: process criteria and product criteria (Guskey, 1994). Teachers utilizing process criteria are considering the development of learning and growth within a program. They report the results of quizzes, homework, class participation, or even attendance. They are also cognizant of effort and work habits. On the other hand, teachers who utilize product criteria base their results on final exam scores, overall assessments based on product appearance, and other critical presentations that demonstrate learning. For these teachers, the significance of where their students are is more meaningful than how far they have come. Process criteria appear to be reliable measures of student growth and establish a baseline from which growth is gauged. Yet product criteria also can provide a valid estimate of student learning as an indication of current ability and knowledge.

Put simply, criteria are statements describing elements of a lesson that are significant in determining whether or not the lesson objectives are being met. It goes without saying that if the lesson is to teach color harmonies, the assessment criteria should be specific to aspects of color harmony. A list of criteria would include proper identification of a particular harmony on a color wheel—the student properly identifies a triadic color harmony. Or when viewing art prints in the classroom, the student properly identifies the color harmony of a given artwork.

Assessment criteria can extend to any and all aspects of each lesson. Another example could be a lesson objective that states, "The student will demonstrate proper use of the vanishing point in a one-point perspective drawing." In this case, the product assessment criteria would specifically designate the proper use of the vanishing point as described in the objective. The design and development of assessment criteria are as vital as that of the design and development of each lesson plan and its objectives (Gruber, 1994). Assessment criteria should be included in each lesson plan.

Balanced Assessment

No single aspect of assessment can provide a representative and accurate measure of student learning in art. Recognizing this, many art education writers call for a variety of assessment strategies that include testing, observation, products, and portfolios. Additionally, there is support for the use of a "balanced approach" that spreads the assessment findings over a wide range of activities (Gruber, 1994; Gruber & Hobbs, 2002).

Testing, observation, finished product, and portfolios, when used individually, do not provide a completely accurate representation of learning in art. However, like the four legs of a table, the four assessment strategies form a balanced support for a comprehensive assessment plan. Because of the diversity of assessment strategies available for each lesson, a more equitable and accurate measure of student learning is possible with a balanced assessment approach that utilizes all four strategies.

Written Tests

The most widely used source of assessment information apart from the final product for art teachers is written tests (Gruber, 1998). Art educator and researcher Karen Hamblen warns us to use caution with tests (1988). "Although testing may appear to be justifiable as a means to give art instruction credibility, it is also necessary to assess whether testing itself may change the character of art instruction" (Hamblen, 1988, p. 60). One shortcoming of written tests is their inability to measure aesthetic responses (Gaitskell & Hurwitz, 1970).

Still, tests can be used as an important aspect of a total assessment approach as long as they remain applicable to the goals and objectives of the lesson (Gruber, 1994). Tests should be used primarily within concrete applications that measure cognitive learnings derived from the subject. These learnings can range from simple recall of names of artists to correctly identifying stylistic characteristics of artworks.

There are basically two types of tests available for use by educators in all subject areas: forced-choice and performance based tests. Forced-choice tests are simple and generally reliable, although their ability to measure objectivity can be limited. They offer students a set of answers from which to choose when responding to test questions or statements. Examples of forced-choice tests are matching, true-false, fill-in-the-blank, and completion. Although reliable in their ability to produce consistent results, they cannot measure student performance. Forced-choice tests are generally too simple to measure higher order thinking skills, and they offer the opportunity for guessing. What's more, test scores derived from forced-choice tests are not valid indicators of much more than students' ability to memorize and restate data. They should not be relied upon as the sole source of assessment for student learning. On the other hand, forced-choice tests are a valuable component of the art educator's assessment tool kit. They are cost effective, can be developed and administered in a short period of time, and can be used as a broad measure of student knowledge. As such they represent one leg of a balanced assessment approach.

There are basically two types of tests available for use by educators in all subject areas: forced-choice and performance based tests.

Observation

Observation strategies involve consistent surveillance of what students actually do during all phases of an art lesson: performance, attitude, work habits, and behavior in general. Criteria developed in the form of checklists or rubrics are used to record observation data. Such data can be aimed at behaviors during the ongoing process of studio activities such as organizing the work area, use of tools and materials, and clean up.

Observational strategies can also take into account behavioral activities of students as they engage in discussion or general class-room decorum. Finally, an appropriate approach "would be to note a student's participation, the level of serious engagement, the sharing of discoveries … and attention to the task" (Day, 1985, p. 236).

Checklists and Rubrics

Observational assessment strategies utilize checklists or rubrics. Checklists are simple lists of contents or attributes against which student behaviors, or projects, products, or portfolios are compared. There is essentially no limit to the length or number of attributes one includes on a checklist. They can be quantified for simple conversion to numerical equivalents or letter grades. For example, a student product that includes 7 of 10 listed attributes converts to 70%. The data can be applied to the total assessment program. Checklists are considerably less detailed than rubrics.

Rubrics, on the other hand, are detailed guides for scoring student behaviors, products, projects, or portfolios. They are specific descriptors listing criteria for levels of expectations or accomplishments from highest to lowest. Upper level accomplishments score higher. Typically, a rubric will consist of four or five descriptors ranging from 1 to 5 points and generally correspond to 4- or 5-point grading scales. In contrast to checklists, rubrics are more difficult to develop, but are applicable to more objective analyses of learning.

Finished Product

The finished product is the result of a lesson and represents one piece of evidence for the measurement of learning. Some art educators claim that focusing on the product will divert students' attention away from the creative process. But even if this were true, it would not prevent the teacher from using the product as evidence to draw inferences about students' creativity and learning in art. Using finished products as a measurement strategy should not lead to using them exclusively in the assessment of learning. They can be directly related to a craftsmanship component by observing, for example, the smoothness of cut edges when students are using scissors, the sureness of brush strokes when painting, or the general appearance of the finished work.

Referring to the quality of student products, education curriculum theorist Grant Wiggins recommends a system of reporting where "degree of difficulty" and "quality" are separated in a method similar to those used in music and gymnastics (Wiggins, 1994). It can be argued that for "any reporting system (to) be complete, it must place the student's performance (product) on a continuum" (p. 35) thus providing some evidence of progress. Furthermore, it should be pointed out that the use of products requires distinguishing between achievement and aptitude (Cramond, 1994).

Discipline-Based Art Education (DBAE) pioneer Michael Day writes that the use of products is a valuable assessment tool because they reflect learning and the results can be pointed out and discussed (Day, 1985). What's more, art education writer Elizabeth C. Clarke points out that products can demonstrate representational accuracy, detail, and expressive quality (Clarke, 1979). Whatever the system or criteria used, the final product remains a significant part of the total assessment mix.

Observation data can focus on the ongoing process of studio activities such as organizing the work area, use of tools and materials. (Student Kylee Brown)
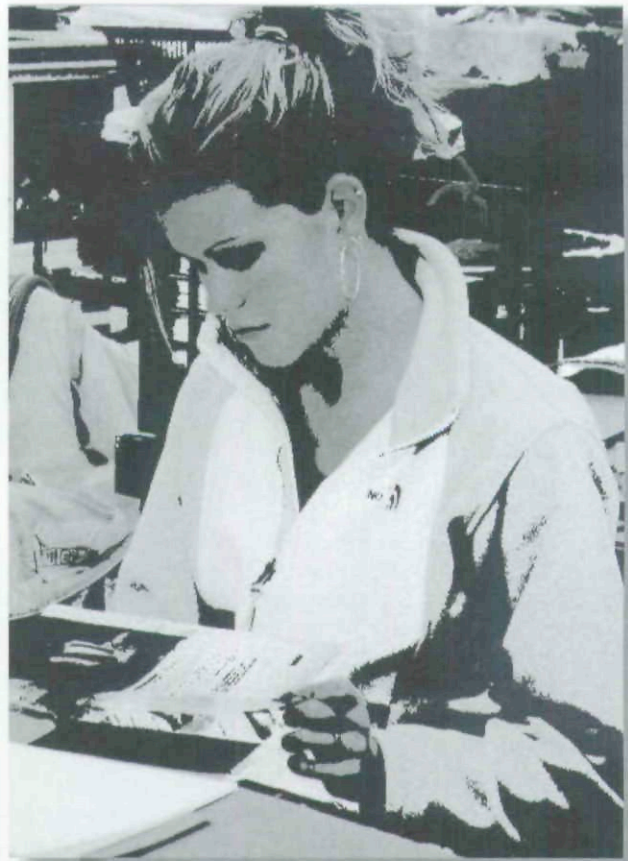
Product assessment can consider the craftsmanship component by observing, for example, the smoothness of cut edges when students are using scissors. (Student Nichole Murphy)
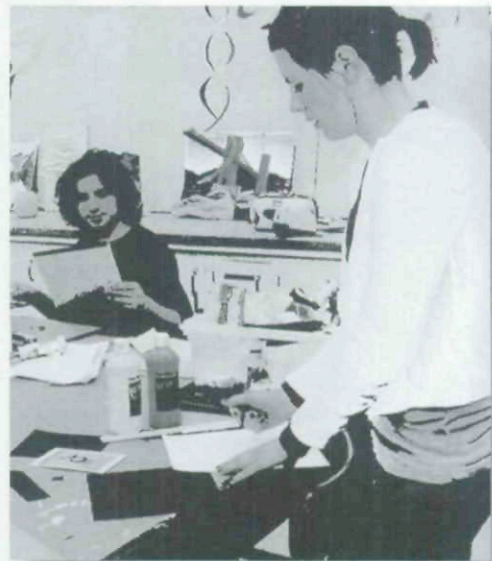
Assessment is an ongoing and post-production activity. (Students Molly Pearce, standing, and Robert Nosek)

Any assessment strategy must consider the objectives of the lesson. It is unproductive to haphazardly assess without regard to the stated objectives. Assessment strategies measure how well the objectives of each lesson are being met.

Portfolios

The usefulness of portfolio assessment is evidenced by its growing popularity across the educational spectrum. Other subject areas are beginning to use portfolios as an alternative to traditional paper and pencil tests. In many respects, portfolios involve the three assessment strategies previously discussed: written tests, finished products, and observation (inasmuch as journals—that is, students' self-observations—may be included). A portfolio is a collection of student work and documents that, as a totality, constitutes a valuable assessment tool in an art program. Portfolios generally fall into two categories: formative and summative.

Formative portfolios contain all aspects of a particular lesson from initial conceptualizations through research and revisions to final products. This information can be in the form of notes, sketches and other artwork, written reports, tests, and any material relevant to the objectives of a lesson. Formative portfolios document the level of involvement, learning, and thought processes that have taken place over a period of time. They represent a more comprehensive measure of learning than any individual final product.

Summative portfolios contain collections of finished products that give indications of abilities at a given point in the instruction, usually at the end of a course of study. Portfolios that contain several examples of finished products only give evidence of ability. They fall short of demonstrating a continuous process of learning and growth. They are, however, useful as verification of a student's aptitude with media and processes.

Using portfolios as an assessment tool is appropriate within a balanced approach. Portfolio assessment is significant when used to show what students have accomplished within an art education program.

Developing Assessment Strategies

Assessment strategies must not be so labor intensive or intrusive that they leave little time for teaching, or that de-emphasize the creative process. Assessment is an ongoing and post-production activity. Ongoing assessment involves observation, and to some extent testing, student self-assessments, and even journals. Post-production assessments involve product (performance assessment), portfolios (formative and summative), and testing.

No lesson plan is complete without learning objectives. The objectives describe what the students will learn as a result of completing the lesson. A typical lesson plan will have sections that indicate scope and purpose of the lesson, lesson objectives, materials, tools, and procedure of the lesson. It must also have a section delineating the assessment procedure(s). Any assessment strategy must consider the objectives of the lesson. It is unproductive to haphazardly assess without regard to the stated objectives. Assessment strategies measure how well the objectives of each lesson are being met.
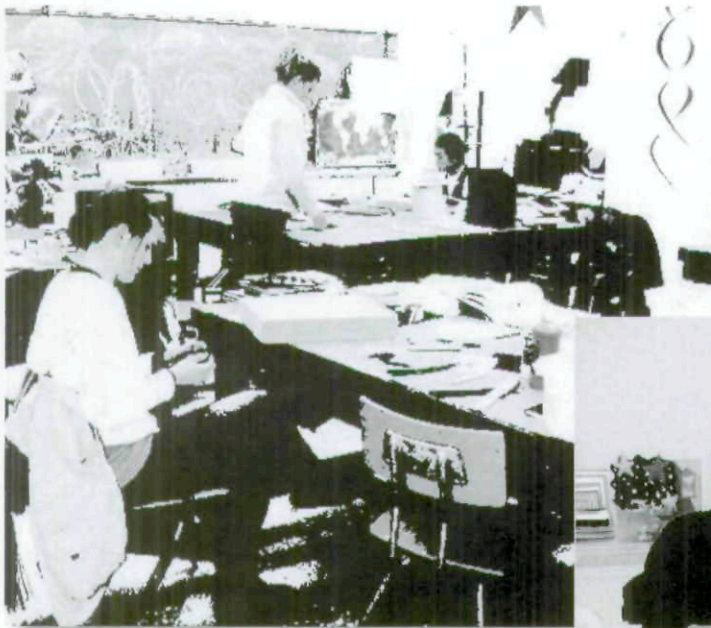
Assessment concerns the learning that takes place within individual students as a result of their interactions within an art program. (Students Molly, Kylee, and Andy)

No single aspect of assessment can provide a representative and accurate measure of student learning in art. (Students Adrianna Jones and Robert Nosek)

Noted Professor of Art Education June King McFee insisted that no matter what assessment strategies are used, it is imperative that teachers possess a set of standards against which to assess each individual product, test, observation, or portfolio (McFee, 1961). Other considerations are students' understanding of art and the ability to respond emotionally to art. Wiggins (1994) agreed with McFee in stating that it is necessary to develop assessment criteria that can summarize and record student achievement.

Donald D. Gruber is Adjunct Professor of Education, Millikin University, Decatur, Illinois. E-mail: teacher1@bwsys.net

REFERENCES

Barkan, M. (1962). Transition in art education: Changing conceptions of curriculum content and teaching. *Art Education, 15*(7), 12-18.

Clarke, E.C. (1979). The educational evaluation of children's artistic progress: More is better, or is it? *Art Education, 32*(1), 20-26.

Cramond, B. (1994). We can trust creativity tests. *Educational Leadership, 52*(2), 70-71.

Day, M. D. (1985). Evaluating student achievement in discipline-based art programs. *Studies in Art Education, 26*(4), 232-240.

Efland, A. D. (1990). *A history of art education: Intellectual and social currents in teaching the visual arts.* New York: Teachers College Press.

Eisner, E. W. (1966). Evaluating children's art. In E.W. Eisner & D.W. Ecker (Eds.), *Readings in art education* (pp. 384-388). Waltham, MA: Blaisdell Publishing.

Gaitskell, C. D. (1958). *Children and their art: Methods for the elementary school.* New York: Harcourt, Brace.

Gaitskell, C. D., & Hurwitz, A. (1970). *Children and their art: Methods for the elementary school,* (2nd ed.). New York: Harcourt, Brace & World.

Gentile, J.R., & Murnyack, N.C. (1989). How shall students be graded in discipline-based art education? *Art Education, 24*(6), 33-41.

Goodenough, F. L. (1926). *Measurement of intelligence by drawing.* Yonkers, NY: World Book Company.

Gruber, D. D. (1994). Evaluation of student achievement in K-12 art education programs: A balanced approach. *Art Education, 47*(2), 40-45.

Gruber, D. D. (1998). *A survey of assessment procedures, uses and perceptions among K-12 art teachers.* UMI Dissertation Services, Ann Arbor, MI.

Gruber, D. D., & Hobbs, J.A. (2002). Historical analysis of assessment in art education. *Art Education, 55*(6), 12-17.

Guskey, T. R. (1994). Making the grade: what benefits students? *Educational Leadership, 52*(2), 14-19.

Hamblen, K. A. (1988). If it is to be tested, it will be taught: A rationale worthy of examination. *Art Education, 41*(5), 59-62.

Hardiman, G.W., & Zernich, T. (1981). Curriculum considerations for art education. In G.W. Hardiman, & T. Zernich (Eds.), *Foundations for curriculum development and evaluation in art education* (pp. 191-205). Champaign, IL: Stipes Publishing.

Lowenfeld, V. & Brittain, W. L. (1975). *Creative and mental growth: A textbook on art education* (6th ed.). New York: MacMillan.

McFee, J.K. (1961). *Preparation for art.* Belmont, CA: Wadsworth Publishing Co.

Meier, N. C. (1966). Factors in artistic aptitude: final summary of a ten-year study of a special ability. In E. W. Eisner & D. W. Ecker (Eds.), *Readings in art education* (pp. 105-116).

Rush, J.C.,(1987). Interlocking images: The conceptual core of a discipline-based art lesson. *Studies in Art Education, 28*(4), 206-220.

Rush, J. C. (1989). Coaching by conceptual focus: Problems, solutions, and tutored images. *Studies in Art Education, 31*(1), 46-57.

Smith, R. A. (1966). *Aesthetics and criticism in art education.* Chicago: Rand McNally.

Wiggins, G. (1994). Toward better report cards. *Educational Leadership, 52*(2), 28-37.

AUTHOR'S NOTE

All photos by the author, used with permission.