Assessing Art Learning in Changing Contexts: High-Stakes Accountability, International Standards and Changing Conceptions of Artistic Development

Doug Boughton
Northern Illinois University

INTRODUCTION

Underpinning art education since the 1940s has been the assumption that student autonomy is not only important but also central to their art making. It has been often argued that student experience in school art programs develops the capacity for independent thought and the ability to express ideas in visual form. Such individual expression has long been valued in American schools. Recent educational reforms have placed significant emphasis on testing as a way to improve standards in school subjects across the board in the United States. An unfortunate byproduct of these reforms has been the homogenization of student outcomes expressed as standards that are frequently measured with inappropriate assessment instruments. High-stakes tests, employed by state assessment authorities, provide the model for inappropriate assessment. These tests imply homogenous outcomes reflecting a single set of agreed standards is appropriate for the arts. The casualties in these reforms have been the most valued of all tenets of art education: the freedom of students to pursue independent learning pathways and the autonomy of their expression.

This state of affairs is more than an assessment problem; it is a complex curriculum issue. The intimacy of the relationship between assessment and curriculum is the focus of the discussion in this chapter. Here I will focus on three main themes that address the relationship between curriculum and assessment. The first is the importance of supporting student autonomy in their development of artistry; the second is to represent the ways in which the international community does this, with specific reference to the International Baccalaureate Organization; and the third is to provide a view of the potential use of technology in assessment to support a postmodern conception of artistic development that enables multiple pathways for students.
HIGH STAKES ACCOUNTABILITY AND STUDENTS’ AUTONOMY AS ARTISTS

Until recently the issue of student assessment in art education received scant attention in the United States. Between 1959 and 1974, for example, only 5% (14) of the articles published in *Studies in Art Education* dealt in any way with evaluation issues and these were concerned mainly with program evaluation rather than student assessment. However, current accountability pressures in the United States have caused arts educators to pay increased attention to the difficult business of assessing learning in the arts. High-stakes assessment in particular has heightened schools’ awareness of the need to assess student learning in all subjects with an unprecedented intensity.

The arts, although currently subject to high-stakes assessment in only a few states have nevertheless been caught up in the backwash of the assessment momentum created by individual state accountability efforts. Arts teachers in the United States do not have the same history as their European counterparts in large-scale arts assessments, and are not as well equipped to deal with demands to demonstrate publicly the quality of their students’ achievements through regularized assessments of their subjects. American teachers are much more accustomed to making individual classroom judgments about their students’ learning that are not challenged by external reviewers or compared to the judgments made by other teachers in other schools. A large-scale survey of art teachers conducted by Burton found that the majority of art teachers conduct assessment by informal means, “such as observation of students (87.3%), viewing artwork (75%), critiques or evaluations upon completion of projects (62.7%), or conversations (50%)” (Burton, 1998, p. 1). This professional autonomy has long been treasured in the American context so that recent demands for public testing of students in the arts has lifted the curtain on what was previously a private stage for art teachers.

Overseas teachers of the arts are more familiar with public scrutiny. European and Australian countries, in contrast to the United States, admit students to university based on state or national public examinations of senior school subjects (including art), rather than on administering standardized university admission exams. Rigorous summative assessment procedures at the senior school level in art education have a lengthy tradition in Europe, the United Kingdom, Australia, New Zealand, and parts of Asia. Portfolios in various forms play a significant role in these assessments. This is the case, not only for art making but also for art history and design.

Unlike much of the rest of the developed world, high school grades, or aggregate high school scores, do not provide the sole or even the major gateway to universities in the United States. In only three states, Arizona, Kentucky, and Minnesota, do all universities require study of the arts for admission. In nine other states some universities require the arts for university admission. In other words the study of the arts in high school is not required for admission to university in three quarters of the nations’ schools (2002–2003 State Arts Education Policy Database). Students in the United States gain admission largely on the basis of their ACT or SAT Scores. Admission to art schools on the other hand is largely based on portfolios for entry to foundations programs.

Public scrutiny of U.S. education has brought a search for more defensible art assessment procedures. For some insights into some alternative ways of conceiving large-scale assessments I will later look outside the United States to examine the portfolio assessment practices employed by the International Baccalaureate (IB) program at the senior school art level. The procedures employed in this large-scale assessment program, which is taught in over 60 countries, can inform both high stakes testing programs and individual classroom assessment practice.

The impetus for increased attention to assessment and teacher accountability has been growing steadily since the mid-1970s (Davis, 1993). In 1994, the Consortium of National Arts Education Associations published the *National Standards for Arts Education* as a major...
as a contribution toward the effort to improve arts teaching in the United States. It was assumed that these statements about the basic content of the field would provide a baseline reference for teachers across the nation and that assessment practices based on agreed outcome statements would improve the art learning of students. This claim was identical to that being made in other academic fields such as science, mathematics, and social studies . . . despite fundamental differences between the arts and other academic disciplines, and the ways in which excellence is judged in those disciplines. In the same year, the Goals 2000: Educate America Act emphasized the standards to be applied to curricular content, student performance, and opportunity to learn. In the 8 years since that publication, interest in assessment has accelerated in response to heightened accountability measures.

The development of national standards was regarded by many as an historic achievement in the United States (Armstrong, 1996). The majority of the United States has subsequently adopted a version of the national art standards that have formed the basis for state school curriculums and assessment practices. Of greater significance nationally has been the development of standardized high stakes testing designed as an accountability measure to determine which schools have been able to measure up and which have not. The basis for reference in the development of test items is the national standards or the state adaptations of them. Art, however, is rarely included in the basket of subjects used in these high-stakes tests. Only seven states require the arts to be tested statewide (2002–2003 State Arts Education Policy Database).

The majority of those seven states employ paper-and-pencil multiple-choice test formats. Some are combined with other options such as open-ended written responses and portfolios. Arts assessments are also commonly combined with other subjects within the humanities.1 The effect of testing the standards through the use of multiple-choice tests is to atomize artistic knowledge to the point where assessment information gathered is virtually meaningless. Consider the following questions selected from the 1996 Illinois state art test (Illinois State Board of Education, 1996). Students were required to answer 13 art questions in total. No assessment of art performance is undertaken in Illinois.

Q14. Which word best describes the theme for this sculpture?
   A. Unity
   B. Chaos
   C. Reality
   D. Tragedy

Q15. What is a style of writing that is also considered an art form?
   A. Graphic art
   B. Narrative art
   C. Iconography
   D. Calligraphy

Q19. Which art process uses a woodblock?
   A. Printmaking
   B. Painting
   C. Photography
   D. Weaving

Q22. Which painting has a horizon line below eye level?
   A. 1
   B. 2
   C. 3
   D. 4

---

1Information about art assessment in these states was gathered from state education Web sites and follow-up phone calls (October–November, 2002).
Q23. Which painting has a horizon line above eye level?
   A. 1
   B. 2
   C. 3
   D. 4

The national standards for arts education, between kindergarten and 12th grade contain 57 standards for the visual arts. Assessment information, if collected in relation to each one of these standards, provides detail about individual parts that have little intrinsic value. Examination of holistic acts of artistic achievement is complicated because the parts are interrelated and many standards are addressed simultaneously. Artistic performance is organic and cannot be dismembered without serious injury.

It is encouraging to see that the 1997 National Assessment of Educational Progress (NAEP) employed a wider range of assessment techniques than is common in state high-stakes tests. Students were asked to create, analyze, and interpret works of art (U.S. Department of Education, 1999). Although the NAEP assessment did not collect data over time portfolio style, the response to given tasks required students to produce artworks with relatively open-ended outcomes in addition to answering some multiple-choice questions about artworks.

Performance measures, such as portfolio or other “authentic” assessment techniques, however, are rare in statewide assessments of art, most likely due to the high cost of their implementation and the likely perception of policymakers that the arts are not worth the investment because their subjectivity compared to mathematics or science would make assessment unreliable. The result is that art is not tested in 43 states, which diminishes perception of its importance in the curriculum, and in the 7 states in which it is tested paper-and-pencil multiple-choice assessment formats are predominant. These kinds of assessments are not widely supported in the arts education assessment literature.

Thus, the current high-stakes assessment practice in the United States has created a parlous and no win scenario for art education. Art is not seen to be valued sufficiently by administrators in 43 states to warrant inclusion in their high-stakes assessment programs, and in the remaining 7 states testing is carried out using means that are largely inappropriate. In a national educational context in which the results of high-stakes assessments can affect the real estate values in school neighborhoods, the salaries of school administrators, and the professional futures of school faculties, art teachers increasingly feel the need to demonstrate student learning through systematic assessment that is ill suited to measuring the quality of holistic performances.

Portfolios as an Alternative to Paper-and-Pencil Tests

Questions about the appropriateness of testing as a means to determine artistic understanding have been raised by others (Gardner, 1996; Sullivan, 1993; Zimmerman, 1994) primarily because of the homogenizing effects of the testing process. Standardized testing using a series of discrete and unrelated items requires students to perform in ways that do not typify the kinds of behaviors that even young artists would use to display their knowledge in the broader social context; and in that sense, testing is an “inauthentic” way to determine whether someone possess specific kinds of cultural and practical knowledge in the arts. Other shortcomings of multiple-choice tests have been widely reported in the literature. They often require only lower order thinking skills; they fail to assess all the important and desirable educational outcomes; they can encourage teaching to the test; they can be used and interpreted improperly (Cizek, 1993).

Authentic assessment strategies, on the other hand, engage students in long-term tasks and meaningful projects that are challenging, complex, and reflect real-life situations (Gardner, 1996; Wolfe, 1988; Zimmerman, 1994). Assessment in the arts should not be conceived as
information retrieval (Sullivan, 1993) but as a means to chart students’ intellectual pathways. Insight into students’ thinking and understanding is not always provided by end products. The Arts PROPEL project conducted in Pittsburg public schools (Gardner, 1989, 1996; Wolfe, 1988) was a long-term research effort that explored, among other things, alternative (authentic) assessment strategies based on portfolios, written reflective material, and dialog (Blaikie, 1994).

The Goals 2000 Educate America Act (1994) emphasized the intention of the U.S. education reform movement that advocated a preference for performance-based assessment as the most appropriate means to assess students’ proficiency (Castiglione, 1996). The testing industry in the early 1990s invested considerable energy into retooling its instruments to meet the demand (Cizek, 1993). By 1996, at least 40 states were either exploring alternative testing procedures or were in the process of changing the process of student and professional performance (Castiglione, 1996). However, since the mid-1990s, the early enthusiasm for performance testing has declined significantly. For example, the California state testing program in the early 1990s supported locally developed exams that contained a high percentage of performance tasks. In 2002, the state’s accountability index makes exclusive use of SAT9, a standardized, commercially produced, paper-and-pencil test.

Curiously, the idea of portfolio assessment, which is historically rooted in the visual arts, has, at the same time, been embraced as a viable solution to the shortcomings of paper-and-pencil testing in other subject areas and is largely ignored both by the state high-stakes assessment programs and by the teachers of the arts in the United States. Burton (1998) found that only 17.1% of teachers use portfolio review as a primary method of assessment.

Authentic Assessment and Portfolio Evidence

Portfolio evidence is one of the most important elements of alternative (authentic) assessment (Gitomer, Grosh, & Price, 1992). Both practice and research indicate that students may usefully collect within their portfolios work in progress; completed works; sketches and notes about ideas related to the work; assessments and commentaries by the student, teacher, and peers; essays about the work; photographs and other records of source materials. Portfolios have been employed widely and successfully for assessment in the visual arts by the Advanced Placement, International Baccalaureate, and Arts PROPEL programs (Blaikie, 1994), the Dutch National Assessment System (Schönau, 1996), the British GCSE art assessment (MacGregor, 1990), and Australian and New Zealand Year 12 state and national assessments (MacGregor, 1990, 1991).

CHANGING CONCEPTIONS ABOUT THE CONTENT OF THE FIELD

In the visual arts, the business of conducting large-scale assessments remains a vexed issue. The modernist philosophy underpinning the approach to assessment by standards assumes that artistic activity and its products can be deconstructed into discrete components, each of which can be assessed individually. The relationship between the components is not addressed by the standards. Assessment becomes a mechanistic gathering of bits of information that, taken one by one, contain no inherent value. In addition the organic nature of the arts is ignored; the relationship among history, culture, context, and production becomes disconnected; and complexity, which characterizes performance in the arts, is lost. In the end it is not the important and complex outcomes that are assessed but those individual pieces that are easy to measure.

In the modernist framework of standards-based assessment, life is oversimplified for teachers in comparison to the tangle of issues raised by changing conceptions of the field of art. The
demands by politicians and school systems managers for art teachers to neatly characterize the nature and content of the field in terms of performance standards has occurred at a moment in the history of art and education which is characterized by vigorous challenges to established modernist orthodoxies. Specifically, new approaches to art education, broadly categorized under the term *visual culture* question long-standing assumptions about both the nature of art and the content and practice of art education.

The visual culture approach is sufficiently well established in the literature to indicate the emergence of a significant challenge to established modernist art education orthodoxies (Efland, Freedman, & Stuhr, 1996; MacGregor, 1992b). Special issues of *Studies in Art Education* and *Art Education* in 2003 highlight this debate. Based on postmodern philosophy, visual culture approaches have proposed curriculum ideas that carry significant implications for assessment. Important among these are: (a) The meaning of visual images is as important as its form; (b) art reflects social issues such as equity, identity, and community; (c) the value of artworks is attributed by the viewer; (d) artworks are valued for the associative power of their symbols and social relevance in specific contexts; (e) art is best understood through sociocultural forms of analysis; (f) no distinction is made between fine and popular art forms with regard to their educational value; (g) contemporary art is relevant to students, (h) recycled imagery is appropriate in the context of postmodern art making; and (i) students are encouraged to take responsibility for their learning under the guidance of a teacher who initiates experiences in the full range of visual culture—including, among other things, comic books, video documentaries, film, computer games, and installations (Boughton et al., 2002).

Differences in views held by those who favor a modernist approach and by others holding postmodern perspectives continue to raise questions for educators about central pedagogical and assessment issues like the virtue of originality in image making (Freedman, 1994b; MacGregor, 1992a), the relationship of art and context (Chalmers, 1985; Blandy & Hoffman, 1993), the relative merits of popular art and fine arts (Duncum, 1996), the influence of new technologies on expression (Freedman & Relan, 1992), the relative value of Eurocentric forms of expression for students of diverse cultural backgrounds (Anderson, 1995; Garber, 1995; Stuhr, 1994), and gender issues in art and education (Freedman, 1994a; Sandell, 1991). Despite vigorous defense of traditional aesthetic values (Eisner, 2001; Feldman, 1992; Smith, 1992), the dust raised by the battle between the proponents of old and new ideologies has challenged the universality of traditional formalist criteria, such as the central role of the elements and principles of design, making it more difficult for teacher/evaluators in the United States to determine the parameters for assessing the quality of their students’ artistic products.

The conceptions of content of the field in a visual culture approach to art education not only have changed, but also the understanding of the ways in which students progress through it have also changed, further complicating the assessment task for teachers. Traditional views of artistic growth presume the student enters the art program “from a naïve, uninformed state and exits from the program in a sophisticated state as a result of teacher interventions and educational activities…” (Clark & Zimmerman, 1978).

New approaches to art education assume students already possess sophisticated knowledge of the visual arts defined more broadly as visual culture. Students’ capacities to understand the visual arts should not be underestimated, given the complexity of the computer games, feature-length films, and television commercials they experience every day. In most important ways, these forms of popular visual culture are as complex as fine art. As a result of their rich visual experiences, outside of school students today are making connections on their own among the diverse forms of visual culture they encounter daily and constructing a range of new knowledge through and about the visual arts.

The curriculum in this context is a creative process and a conceptual space in which students develop their ideas with the aid of teachers who act as critical partners. Curriculum is a form
of mediation between and among students, teachers, and a wide range of texts and images from inside and outside of school. Assessment of student learning must reject homogeneity of outcomes and embrace the notion of individual student pathways. The IB program exemplifies this approach to curriculum, and its assessment provides the right environment for it to work. High-stakes assessment models in the United States do the exact opposite.

**ASSESSMENT IN THE INTERNATIONAL CONTEXT**

To this point I have argued that the current climate of accountability in the United States has raised the specter of public scrutiny of student learning in the visual arts, and that many arts teachers do not have a history that prepares them to respond to this demand well. I have argued also that despite the development of national and state-based standards for the teaching of art, current state testing formats are largely inappropriate to assess arts learning; and further, the field is undergoing a transition that points to a need for a more reflexive, holistic approach to both the construction of curriculum and the conduct of assessment. An implication from the aforementioned is that the relationship between curriculum and assessment needs to be reconsidered in the United States.

Looking more globally at the issues of assessment, the IB program is one that well reflects the challenge of achieving universal excellence variously expressed in multiple personal and cultural contexts. The IB is a system of international education taught in more than half the countries in the world. As described in documents accessed through the information page of the IB Web site, the centerpiece of the program is its flexibility in responding to local interests but at the same time providing access for students to what is shared and what is different in human experience.

...developing citizens of the world-culture, language and learning to live together, building, and reinforcing students’ sense of identity and cultural awareness fostering students’ recognition and development of universal human values. (International Baccalaureate Organization [IBO], 2003a, p. 7)

Probably the most fundamental difference between the IB approach to assessment and the prevailing culture of assessment in the United States is summed up in the following:

... in all such deliberations the primary consideration is always a recognition of the effect that any assessment structure has on classroom teaching. It is essential that assessment structures support good classroom practice. This is considered of greater educational importance than pure reliability of measurement. (International Baccalaureate Organization, 2003b, p. 15)

It is hard to imagine a greater possible difference in assessment cultures. Brent Wilson (1996), after spending 3 months in the Netherlands, England, Scotland, and Wales studying those countries national assessments, arrived at the conclusion that the United States could benefit by taking note of assessment practices in Europe and in England. He found that the public examinations in those countries “created high-level national expectations, and at the same time encouraged teachers to review the effectiveness of their own instructional practices, to collaborate with their colleagues to improve the general quality of art education, and to openly discuss issues relating to student performance” (p. 1).

To be fair, it is important to note that centralized assessment practices in the IB, Europe, Britain, and Australasia experience some important problems related to examiner training, the establishment of benchmarks, moderation systems, and the sheer volume of assessment
materials handled. On balance, however, the portfolio-based moderated assessment systems used internationally are far more supportive of good classroom practice.

The IB program is one that uses both studio and research portfolios as a central assessment tools. A feature of the assessment is that few criteria are employed, and holistic judgments are employed to assess studio art learning.

THE INTERNATIONAL BACCALAUREATE PROGRAM

The IB program has a long tradition of public art portfolio examination in more that 60 countries, including the United States, where more than half the total world candidature is located. In the course of its history, the IB has developed examination procedures in art (and all other diploma subjects) that are accepted by universities in more than 110 countries (IBO, 2001). In the United States, 123 three universities accepted the IB diploma subjects for entry in 2001 including high status institutions such as Harvard and Stanford (International Baccalaureate North America, 2001). The International Baccalaureate Organization, IBO, was founded in 1968 as a nonprofit educational foundation based in Geneva, Switzerland. It has its origins as early as 1924, at which time international schools in Europe attempted to establish a common curriculum and university entry credentials. Thus, it had to develop a truly international program that would resonate in very different educational systems around the world. Critical thinking and exposure to a variety of points of view are believed to encourage not only intellectual and artistic, understanding but also intercultural understanding by IB students. The IB diploma is taken by students in the last 2 years of school before university studies, and the curriculum leads to what is now called a “baccalaureate,” administered in 109 countries and recognized by universities everywhere. Development of the diploma program was made possible by grants from UNESCO, the Twentieth Century Fund, the Ford Foundation, and other groups (International Baccalaureate Organization, 2002b).

The IB Curriculum

The diploma curriculum model is traditionally displayed in the shape of a hexagon with six academic areas surrounding a core. At the core of the program are three distinct components, a subject called “Theory of Knowledge,” an extended essay, and a creativity/action/service component. The core components complement a traditional liberal arts curriculum in which the humanities and sciences are studied concurrently. The core subjects provide a critical role in linking the other content areas together, so that the arts become an integral part of education and not a repository of specialized bits of information. Knowledge in the arts is promoted as a way of knowing to which all humans should have access.

The diploma candidates are required to select at least one subject from each of the six subject groups. These are (a) Language A1, (b) Second Language, (c) Individuals and Societies, (d) Experimental Sciences, (e) Mathematics and Computer Sciences, and (f) the Arts. At least three and not more than four subjects are taken at a higher level (HL); and the others, at standard level (SL). A HL subject requires 240 teaching hours, whereas SL subjects are taught in 150 hours. This enables students to arrange their work to explore some subjects in depth and others more broadly over a 2-year period. Flexibility in choosing subject configurations allows the students to pursue areas of personal interest and to meet special requirements for university entrance (International Baccalaureate North America, 2001). The core subject,

---

2Art is offered in over 60 countries as an option within the IB diploma.
theory of knowledge, is one that informs all other subjects within the program and plays a vital role in developing critical perspectives required throughout the entire diploma.

The Visual Arts Program

Visual Arts is one of three options available within the arts subject grouping of the diploma. The others are Music and Theater Arts. Central to this program is the requirement for students to develop a critical perspective of visual arts within a variety of cultures and to pursue an independent pathway of learning through research and artistic practice. Each of the higher level and standard-level programs require students to complete two components: (a) studio work and (b) a research workbook. At the higher level, the studio component comprises a 70% of the total grade with 10% of that total grade being allocated to the determination of overall growth. The research workbook is allocated the remaining 30% of the total grade with 10% of that proportion allocated to a determination of the degree of integration between the research workbook and the studio work. The same weighting is applied to the Standard Level A program. The Standard Level B program reverses the proportions with 70% weighting allocated to the research workbook and 30% to studio exploration.

A distinctive feature of the entire IB program is the expectation that students take responsibility for defining their own learning pathways. Independence of research is a central expectation for both research and studio components. Imagination is integral to the studio components of both the HL and Standard Level A programs. Little is specified in terms of subject content, so that maximum flexibility is available to students to explore the visual arts in the context of various world regions and across different genres of art, which include a nonhierarchical examination of fine arts, popular arts, tribal arts, and design arts. The program takes pains to avoid the imposition of Western cultural hegemony on students who may be studying the subject in any of 60 or more countries. Two of the fundamental requirements are that students demonstrate knowledge of more than one cultural context, and that their studio work and research workbooks be closely related in terms of the investigations undertaken.

The role of the teacher in this context is to provide impetus for students, teach them how to conduct personal research, undertake analysis of context, engage in critical writing; provide resources, and work as a learning associate with the student as they progress through the program. The relationship of assessment strategies with the curriculum is one that is unique and serves as a useful model for comparison with existing practices in the United States.

Assessment in the IB Program

The learning pathway of students is, to large extent, defined by the assessment criteria used to evaluate both the research workbook and studio components. In this way the relationship between assessment and curriculum is intimate, yet the curriculum does not specify the content of study in advance. The content of programs is individual and negotiated between teacher and individual students. This is not to say instruction does not occur. Programs take shape in the context of student interest and instruction defined by the scope of the teachers’ expertise and judgment about what is appropriate to the interest and abilities of students. Various imaginative approaches to instruction are taken in different schools. For example, Atlantic United World College in Cardiff, Wales, has used instructional salary flexibly to employ a permanent art teacher and temporary artists in residence on a term by term basis to most effectively service the specialized interests of students.

The International Baccalaureate Organization Visual Arts syllabus does not have units or modules from which the teacher constructs a course of study, but instead it provides a framework which
allows teachers to choose a content and activities appropriate to their own and their students’ interests and experience. When constructing the course of study the teacher is expected to bear in mind the visual arts assessment criteria and the specific requirements for the assessment tasks explained in this guide. (International Baccalaureate Organization, 2000, p. 7)

Choice of media and techniques are also negotiated.

Work in the studio may combine several techniques and any medium may be used. Artistic understanding and expression may be taught through various techniques from painting to puppetry, calligraphy to computer graphics, and sculpture to conceptual art. Students may demonstrate mastery in various ways, provided their course of study includes an introduction to arts concepts and techniques.

. . . Priority should not necessarily be given to drawing and painting. Design, for example, maybe as worthwhile an experience as observational drawing. A student may achieve a high degree of sensitivity and skill in, for example, photography, photography, ceramics, or the use of electronic media, without being able to draw or paint well. (International Baccalaureate Organization, 2000, p. 9)

The IB program is regularly reviewed by a committee constituted by the Curriculum Office in Cardiff. The most recent curriculum review of the visual arts conducted in 1997 to 1998, and chaired by this writer, redefined both the assessment criteria and the manner in which judgments are made about student learning to reflect a postmodern construction of the discipline and to remove Western cultural bias from the assessment criteria. In the previous assessment plan, weighted criteria were employed to assess studio work. Imaginative and Creative Thinking and Expression was weighted at 35%, Persistence in Research at 20%; Technical Skill at 15%, Understanding the Characteristics and Function of the Chosen Media at 10%, Understanding the Fundamentals of Design at 10%, and Ability to Evaluate Own Growth and Development at 10%. An analytical scoring system was used requiring separate marks to be assigned by examiners against each criterion and the previously cited weightings were then applied to arrive at the final studio grade. Such a process was viewed by the review committee to be inappropriate because it applied an inflexible modernistic construction of art to the assessment of culturally varied student work and promoted homogeneity of outcomes based on Western fine art traditions. Application of a single conception of excellence was contrary to the expressed philosophy of the IB program so the analytical scoring system was replaced by a holistic assessment of student work. In this model, criteria are used for reference, but examiners are free to pay attention to the work in the way that is most appropriate to its genre.

The specifics of these changes are discussed in the following section.

Assessment of the Studio Component

Students at the Higher Level and Standard Level A are required to exhibit a portfolio of their work containing both finished artworks and working pieces demonstrating the process of their visual research. An external examiner visits the school to interview each candidate for 30 to 35 minutes about their work. The examiner allocates a single holistic grade after the interview and following consideration of the work in relation to the following five criteria:

• Imaginative Expression. At the highest level of achievement, the candidate’s explorations are creative and imaginative. Ideas and forms are consistently and intelligently presented in an adventurous manner, resulting in surprising and unusual images, which challenge existing conventions. Unusual combinations of forms, techniques, and media and/or combinations of form and content are frequently evident.
Purposeful Exploration. At the highest level of achievement, there is evidence that the candidate’s explorations of ideas are clearly and strongly integrated with his or her life and cultural context. The candidate includes both analysis and synthesis in the investigations, resulting in a powerful and significant body of work.

Meaning and Function. At the highest level of achievement, the studio work exhibits a synthesis of conceptual content, formal knowledge, and technical skill. It has strong personal, sociocultural, or aesthetic meaning. The relationship between form, function, and meaning is very clear and appropriate.

Formal Qualities. At the highest level of achievement, the studio work consistently shows strong evidence of a thoughtful and inventive use of elements and principles of design. This has resulted in the production of strongly unified works. A comprehensive ability to solve formal and technical problems is clearly evident as demonstrated by rigorous investigation of aspects of form in the body of work.

Technical and Media Skills. At the highest level of achievement, the studio work shows an outstanding technical confidence and demonstrates a highly appropriate use of media in relation to the intended expressive purposes of the work (International Baccalaureate Organization, 2000).

Assessment of the Research Workbook Component

At the Higher Level and Standard Level A, this component is assessed by the teacher in relation to the following four equally weighted criteria using a six-level rubric for each.

- Independent research
- Critical research
- Contextual research
- Visual research

Moderation of Grades

What is moderation? It is a judgment process undertaken by teachers within the community of peers to ensure that the equivalent work done by students in different classrooms and different schools is rated equally. The grades issued by both external examiners and teachers are not the final grade. Moderation is a system of multiple judgments made by different examiners about the students’ work. The intention of moderation is to reduce variations of interpretation among different examiners and serves to promote a climate of debate and discussion about the quality of student work. This debate is essential in assessment context where students are required to push the limits of their own understanding, to take risks, exercise imagination, and interpret the visual world critically. The best students frequently produce work that will perplex examiners, and this is the way it should be if art is properly taught in a postmodern context. A second, and sometimes third look, at student work is often necessary to determine its qualities and to serve students fairly.

This process is particularly important to the international Baccalaureate program given the wide geographic distribution of students who participate in the program. Different examiners are employed to visit schools in over 60 countries in which art is taught as part of the IB Diploma. Subsequent to the examiner visits to schools, photographic and photocopy samples of candidates’ work is sent to a central location in Cardiff, Wales, where a team of experienced and trained moderators, under the direction of a chief examiner, compare the visiting examiners’ and teachers’ judgments against agreed benchmarks of performance. Benchmarks of the best work are drawn from the international student community and posted year by year on the IB virtual gallery accessible via the internet. These works are available for access by
teachers, students, and examiners. Benchmarks illustrating the range of achievement at specific levels from highest to lowest are sampled from student work and made available to examiners only.

The benefit of this process, in addition to ensuring more reliable judgments for students, is that examiners and teachers receive feedback about their judgments, thus developing a community of agreement about standards. The IB is not the only program to employ moderation procedures. Moderation is used on national scale in the United Kingdom (Steers, 1988), the Netherlands (Schönau, 1996), Australia (Boughton, 1994), and by the AP program in the United States (Askin, 1985).

Benchmarking

Central to the moderation process is the practice of benchmarking. In simple terms, benchmarks are samples of student work selected by moderators to exemplify specific levels of achievement. The work samples clearly indicate the limits of performance within each level. If, for example, five levels of performance are specified by performance descriptors, five collections of studio work are selected to define the limits of each level. Written performance descriptors alone tend to be limited in their ability to represent the qualities of visual art. Therefore the benchmarks take the form of actual examples of student work.

Benchmarking is an idea that has been much practiced in U.S. businesses (Codling, 1998) and is now finding favor in higher education (Alsete, 1996; Barak & Kniker, 2002; Tucker, 1996). However, in the United Kingdom and Europe, benchmarking has been practiced for many years in art assessments as well as in other fields. There are many approaches to the selection of benchmark work (Boughton, 1997). It is possible to select benchmarks each year from the cohort of candidates who are to be assessed. It is also possible to choose work from past years to represent benchmark standards. A combination of both past and present work may also be chosen. Irrespective of these choices, the idea is to choose multiple samples of work that represent the lower bound, the center, and the upper levels specified in the system. The visual arts are dynamic and unpredictable, thus, the intention is not to choose examples that must be matched by student candidates’ work. Rather it is to choose samples of work that represent qualities rather than specific models of performance. In other words, an excellent painting of a scene depicting poverty in Indonesia is not intended to provide an image for students to copy in order to receive high grades. The painting is intended to exemplify an imaginative representation of a political statement, superior understanding of media, expressive use of form that is supportive of the content of the work. Students who attempt to make copies of benchmarks are penalized in their assessment.

The IB program provides benchmark of student work to examiners, selected from previous students’ work. Teachers and students are provided with examples of high-level work chosen from previous years’ examinations on the IB Web site (http://online.ibo.org/gallery/).

WHAT CAN BE LEARNED ABOUT ASSESSMENT FROM THE IB PROGRAM?

The foregoing description of the IB program provides a useful reference for analysis of trends in the assessment of student work in the visual arts. Research pertaining to art education assessment is scant, with far greater attention paid to the more attractive issues of curriculum content. As a result, assessment practice of student work in visual arts often draws important ideas from practice rather than from research. The IB program provides a unique model that incorporates best practice based on literature.
Important Defining Characteristics of the Portfolio

Although much has been written about the portfolio, it is an instrument that is frequently misused or misunderstood. The commonly understood characteristic of a portfolio is that it is a collection of work accumulated over time. There are, however, some other common features of traditional portfolios that, if overlooked, reduce the potency of the portfolio as an assessment tool. The first of these is that the content of the portfolio process is embedded in the ongoing program of instruction but open ended in the sense that students are encouraged to develop classroom experiences into independent explorations of ideas (Stecher & Herman, 1997).

This central intention here is that portfolio entries should be derived from regular instructional events and are not be the result of “on-demand” tasks. The student should be free to interpret the ideas encountered both inside and outside class and to develop independence in their exploration of art ideas. This characteristic, if present, enables students to take risks and move beyond classroom exercises.

Taking responsibility for learning and developing the capacity to work independently are important indicators of good art learning, and this is a central characteristic of the IB program. Not only does the portfolio serve as an assessment tool, it also plays a vital role in the meaningful elaboration of curriculum intentions. In short, the portfolio becomes integrated with curriculum in very important ways and is not simply a repository for all class assignments set throughout the year.

Burton’s (1998) study found that 52% of all visual arts teachers surveyed in the United States assessed their students at the completion of each studio project or written assignment. The portfolio that contains only this collection of assigned work and lacks open-ended content is one where the teacher defines both the content and the outcome of each project. Such practice will ultimately proscribe the form and content of the portfolio. At the end of the term, semester, or year, students in the classes of these teachers will typically present portfolios that look very much the same as one another with products that meet the common project criteria demanded by the teacher. These kinds of portfolios do not reflect the student’s capacity to work independently, nor do they reveal the degree to which students are willing to take risks in order to extrapolate from, and interpret, the ideas presented in class. By definition, the only thing these portfolios can do is showcase the teacher’s capacity to invent tasks for student response and to direct their outcomes.

A second feature of good portfolios identified in the literature is that they contain student-selected entries (Castiglione, 1996; Stecher & Herman 1997). Although the idea of educational portfolios are prominent in the professional artworld the educational application of portfolios is different (Castiglione, 1996). The artist portfolio is usually a display of a person’s public professional persona and does not usually contain works indicative of process, doubts, or failed explorations. The purpose of education portfolios is to promote students’ knowledge of their own progress and to support their ability to demonstrate independence in researching and evolving projects of their own. Thus, works in progress, sketches, and reworked pieces are important as portfolio entries because they provide insight into student growth and the pattern of decisions students have made in relation to their evolving work.

Without student choice, there is no indication of the student’s capacity to make informed decisions about his or her own ideas and progress. Often it is possible to discover as much about a student by what they choose to include as it is from the quality of the work itself. Clearly, the degree to which this is possible is determined to some extent by the age and sophistication of the students involved. Less is expected of younger students, whereas more fully resolved work can be anticipated from senior students. Nevertheless, some choice is possible at all levels of schooling. The IB assessment criteria, Purposeful Exploration (Studio), and Independence of Research (Research Workbook) have evolved from the understanding that portfolios effectively reveal these qualities in ways that other assessment instruments cannot.
A third feature of good portfolios is the importance of student critical self-reflection, which may appear in journals or portfolios in written, or taped form (Wolfe, 1988). In addition to the IB program teacher/examiner interviews have also been used in other programs, such Arts PROPEL (Blaikie, 1994). Ross, Radnor, Mitchell; and Bierton (1993) found that, during reflective discussions with students, teachers tend not to listen carefully to students; that they seem to drive their own agendas through teacher talk; that students understand more about their own feeling states and sensibilities than adults comprehend; and that dialog, properly conducted, can reveal valuable insights into the process of art making, particularly students’ understanding of the quality of the work, the manner of its production, the reasons for choices, influences on the work, difficulties encountered, new ideas to explore, and so on. Here is an example from an IB student’s candidate statement submitted as part of the exhibition of studio work:

I found that seeing various artist’s works has influenced me also. An example of this is Leon Kossoff’s work: after seeing it, I immediately felt like painting in thick, bold brushstrokes. (International Baccalaureate student, May 1996)

This is another example from an IB student:

Drawing human figures is one of my strengths. I used to copy human figures from comics. But then I found out that I needed to learn how to draw real human figures, so I started studying realistic human drawings and I learned a lot from these master drawings, especially Michelangelo, who is one of my favorite artists and has been my greatest inspiration. (International Baccalaureate student, May 2001)

Choosing Criteria for Assessment

The virtue of thoughtful and process focused portfolios resides in its capacity to collect assessment information over time. As was discussed earlier, not all portfolios are created equal. Without carefully considered judgment, the portfolio remains simply a repository of data. For portfolios to be effective, the relationship of assessment criteria to program intentions must be carefully articulated. In the case of the IB, assessment criteria are integral to the program, effectively defining expected outcomes of learning, and drawing upon the strengths of the portfolio strategy to reveal student learning. Choice of criteria can bear a strong relationship to the issue of validity in assessment. Examination of the literature indicates that there is a close relationship between IB criteria and those used by visual arts teachers in other programs internationally.

MacGregor’s (1992a) examination of assessment practices in Britain, Holland, Australia, and New Zealand, showed that the broad criteria used by teachers vary little from country to country. Most teachers employ some variation of three categories of criteria: (a) relative ability to develop and interpret a theme, (b) level of technical expertise, and (c) relative ability to achieve sensitive personal expression through the use of a variety of techniques and processes. The particular strength of the portfolio strategy reflected in the IB program is the manner in which students’ pursuit of ideas and themes can be tracked through both research workbook and studio work.

Making Judgments: Holistic Versus Analytic

One of the most difficult stages of the assessment process is making value judgments about student learning and student products. Considerable debate exists in the field about who should make these judgments and how. The IB program’s reliance on external examiners and the use of
moderation have succeeded in attracting widespread confidence in the reliability of judgments, and there is much to be said in support of the value of moderated assessments which are used throughout Europe, the United Kingdom, and Australasia (Boughton, 1996b; Blaikie, 1994; Wilson, 1996). Castiglione (1996) argues in support of the notion that faculty judgments in art instruction “are necessarily expected to involve commonly accepted standards, to transfer across individuals, and to generalize across judges—in short, to be independent of any criteria unique to the assessor.”

In a comparison of the AP, IB, and Arts PROPEL projects, this is exactly what Blaikie (1994) found. All three programs depended on skilled judges with knowledge of the visual arts, using a set of established criteria to arrive at judgments of student work.

Schönau (1996) reported work done by the Dutch Institute for Educational Measurement (CITO) to examine the Dutch Central Practical Examination (CPE) in terms of the potential for achievement of common standards against national prescriptive criteria. In the first year of the exam, student work (1981) was judged by the student’s teacher and a panel of five judges. In 1982 the jury was reduced to three, and from 1983 the work was judged by the students’ own teacher and a specialist colleague from another school. The analysis of scores showed an acceptable level of judge difference between the scores allocated irrespective of the number of outside judges. However, Schönau also found that global judgments tended to produce higher judge agreement than analytical judgments using criterion analysis.

In 2001, the IB program moved from analytical criterion assessment of studio work to holistic judgments. The studio assessment prior to 2001 comprised six weighted criteria valuing imagination at more than twice the weighting given to technical skill. The problem confronted by curriculum developers was that the system of weighted criteria predetermined a western modernist construction of art that was universally applied across all cultures. Such a unilateral application of criteria was not responsive to cultural context or artistic genre. The move to holistic assessment enabled a far more reflective assessment strategy that took into account the complexity and variation of artistic expression without loss of reliability in the judgment process. Different artwork demands different attention; thus, the use of a single holistic judgment enables judges to pay attention to the specific qualities of individual works. Application of standardized weighted criteria to all genres of work does not permit such flexibility.

What Are the Benefits of the Portfolio?

The greatest single benefit of the traditional portfolio is the flexibility it provides to the program for students to explore, take risks, reflect on their progress, and to exercise fully the skills required to perform in the complex ways demanded by the visual arts. A concomitant benefit is the insight it provides teachers and examiners to understand the students’ development over time (Grace & Shores et al., 1994; Stecher & Herman, 1997). If students have the freedom to make choices about the content they include in their portfolios, and are also encouraged to explore ideas independently, outside the limitations of classroom exercises, then a clear picture of their intellectual footsteps is represented in the contents. I have seen this illustrated vividly in the portfolio exhibitions offered by IB students at their final examinations following 2 years of independent studio exploration.

A second benefit offered by the portfolio is the capacity that it can take multiple forms, each with the capacity to support distinctive ways of working (Grace & Shores et al., 1994). For

---

3An extensive research project was undertaken in 1998 to 1999 by this author in partnership with Kerry Freedman and the International Baccalaureate examinations office to compare the differences in reliability between analytical and holistic studio assessment methods. No statistical difference was found between the two methods using experienced examiners from different world regions.
instance, the working portfolio is an archive of works in progress that provides opportunity for teachers and students to reflect together on work done in the past and to revisit ideas and avenues that may have been forgotten or overlooked. The creation of an exhibition portfolio, on the other hand, requires students to select best work for exhibition or examination. In the context of an IB examination, students present their best work as an exhibition and provide the remainder to serve as backup records for reference, or to provide a point of discussion with the examiner or teacher.

A third benefit offered by portfolios is the motivation it provides to the student. The capacity to review work and see improvement can be a great stimulus for student learning. The reflective component of portfolios, if well used, can promote greater involvement by students with their work and also helps the teacher understand what is going on with student learning. Even math teachers have noted the benefits of portfolios. One algebra teacher reported that she came to use more varied kinds of instruction (e.g., more problem solving and more long-term situational problems) because she used portfolios in her algebra class so that her students ended up having a variety of items to choose from in creating their portfolio. This algebra teacher also noted that portfolios gave insight into students’ maturity, self-esteem, writing ability, and their ability to evaluate their own and other students’ work (Knight, 1992, cited in Lester, Lamdin, & Preston, 1997). Two other researchers, Lamdin and Walker (1994), found that students often became much more reflective about their own mathematical performance when they assumed responsibility for preparing a portfolio of their work.

DIGITAL PORTFOLIOS AND THE FUTURE

Given the frequent benefits of portfolio assessment evident in the IB program, it is important to raise the question of the potential of digital portfolio assessment. Faculty and students at Northern Illinois University have engaged for some years in examining this question at elementary, middle school, high school, college, and in-service levels. Stanley Madeja, working for the past 2 years with Charles Dorn, Robert Sabol, and others, has investigated the development of authentic assessment tools in the visual arts as part of a research grant funded by the National Endowment for the Arts (Dorn, Madeja, & Sabol, 2001).

This project involved the development of authentic assessment strategies with groups of art teachers in Florida, Indiana, and Illinois. Ten secondary teachers and one elementary teacher became interested in the application of digital portfolios for their own classrooms (Fitzsimmons, 2003). Two of these teachers, NIU graduate students, have played leadership roles in spearheading the trial of digital portfolios in state schools.

In their work with teachers using digital portfolios, NIU researchers found many benefits that are identified in the literature by people who have examined the same questions in other settings.

Efficiency: The most frequently mentioned benefits of electronic portfolios is the efficient handling and retrieving of information and images related to student work. First, electronic portfolios, like traditional portfolios, show a clear picture of growth over time (Wiedemer, 1998). But the work in traditional portfolios is far more difficult to organize in any desired sequence and takes far longer for the teacher to thumb through in order to analyze the collection. Pieces are more likely to get lost or distributed to other geographic locations. Electronic portfolios, on the other hand, offer the opportunity for work to be viewed in thumbnails all at once, and/or to be quickly organized into chronological or thematic order. In short, the benefit here is efficiency and safety of organization and data storage.

Second, parent–teacher conferences are easy with small, convenient, and accessible documentation (Guhlin, 1999). NIU graduate students/teachers reported that conversation with parents become much easier when electronic files or when hard-copy printouts of student
26. ASSESSING ART LEARNING IN CHANGING CONTEXTS

Portfolios are on the table. Reference is quick, and teachers have found that it is very easy to explain why a student is not getting good grades by comparing their work to benchmark exemplars produced by other students. This can also be done with traditional portfolios, but electronic versions are simply easier to handle. Again the payoff is efficiency.

Third, accessibility to records is enhanced. Fornander (1999) reported the benefits of electronic portfolios particularly accessibility, portability, utility for examination, and the capacity to widely distribute files. Information can be shared easily among colleagues, making dialog about images possible in a more efficient way.

Fourth, storage is easier. Work does not have to be kept in schools. Completed projects can be taken over immediately and not left to wait for grading. This is a particular benefit for three-dimensional work (Tuttle, 1997). In addition, far greater volumes of work can be stored on disk than could possibly be stored in the art room (Oros, Morgenegg, & Finger, 1998).

Motivation

Computers are effective motivational tools. Teachers and researchers alike reported on the enthusiasm demonstrated by students when working in electronic environments. Kerper (2000), an NIU researcher/observer in an electronic portfolio project classroom, reported high levels of enthusiasm demonstrated by elementary students even when working with inadequate technology. Electronic portfolios also seem to encourage ownership, pride, and an increased level of self-esteem, a factor noted by other researchers (Davis, 1999).

Curriculum Benefits

Because the craft of constructing an electronic portfolio becomes so much a part of the curriculum, NIU researchers have noted some additional curriculum effects. One teacher said that the portfolio process put more pressure on her to clarify her explanations and assessment standards. In short, she found it enforced a kind of discipline that was not necessary under traditional art teaching conditions (Fitzsimmons, 2003). Digital portfolios have the potential to stimulate students to further develop their art pieces made with traditional media. Once students have scanned their work, the possibility exists to enhance it, or even redevelop it in digital form.

Communication Benefits

Because the original artwork goes home after it is done, rather than remaining at school, parents are more aware of what is happening in school. Students are more likely to remember the concepts they have learned and not forget with the passage of time (Kerper, 2000).

Another communication benefit of digital portfolios is that students’ work is able to travel with them throughout their career (Fornander, 1999). This has particular application at the college and university level.

Ancillary Benefit

Both students and teachers in schools have noted that electronic portfolios teach computing skills as ancillary benefit. Because the acquisition of computer skills are regarded as a good thing, the collateral learning in the area of digital technology that takes place is valued.

Assessment Gains

Having listed the Previous benefits, it would be easy to forget the major question: Can digital portfolios actually assist teachers to make better judgments about the quality of student
learning? The direct assessment benefits identified both in the literature and by our Illinois teachers were the following:

**Student Self-Assessment**

Learners are able to see the big picture of their own progress more easily. Because of the ready access to their own record of progress, students seem to become more aware of both the quantity and the quality of their own work, particularly at the senior secondary level.

**Teachers’ Oversight of Curriculum**

Program accountability becomes more evident. Teachers are able to keep a more comprehensive record of student work and overall progress of student groups. Comparison with program goals is easier, making diagnostic assessment easier as well.

**Benchmarking**

Electronic portfolios are useful for benchmarking (Guhlin, 1999; Niguidula, 1998) and are also very useful for the development of exit standards (Fornander, 1999). Traditional portfolios also serve this function, but the digital form has far greater utility.

**CONCLUSION**

It is clear that the current assessment climate in United States is not a healthy one for the arts. I have argued that the resilience of standardized tests and reliance on multiple-choice formats for determining the products of art instruction at the state level serve only to perpetuate an atomistic approach to the gathering of disconnected shards of information, which serves a limited purpose. Such testing has more to do with political posturing than it does with serving the interests of students.

In the classroom, teachers who are pressured to demonstrate the outcomes of their programs are not well prepared, nor well served, by the model of state assessment practices. Attempts to demonstrate that students have met national or state goals result in the gathering of disconnected pieces of information that are easy to collect but are not necessarily reflective of holistic engagement in art making and substantive understanding of the complexity of the arts.

The models of practice employed by the IB program and by other European and Australasian national assessment systems provide some insights into more productive ways to judge student learning. These approaches preserve the essential qualities of artistic performance at high levels and take into account the dynamic nature of the discipline of the visual arts across different cultural contexts. The fundamental instrument used to facilitate these assessment strategies is the portfolio. The portfolio on its own, however, is not the complete solution. There are important conditions that proscribe appropriate use of the portfolio that must be observed in order to achieve the good assessment outcomes.

Concomitant with the portfolio as an instrument are other essential elements of assessment that include benchmarking, moderation processes, and the institutionalizing of debate about the qualities of student art making among the community of peers involved in the education process. Admittedly, these are expensive practices to implement on the state level. However, there is no reason why, given the will, adaptations cannot be developed at the local level to achieve improvement in the benefits of assessment for both students and teachers. Some instances of attempts to do this have already been noted (Dorn, 2002).
The advance of technology has provided yet another possibility for the future. The potential of digital portfolios provides an expanded view of the ways in which assessment information may be gathered and managed. Ultimately the decision to deal with the issue of assessment is one that must be taken in the context of curriculum improvement. The unique quality of the IB program lies, not so much in the way in which the portfolio is used for assessment but in the intimate relationship between assessment practices and the curriculum itself.

If the United States wants to develop a more effective and appropriate approach to assessment of the visual arts in education, such as that found in the IB and European settings, there is a long way to go. However, current assessment practices draw on the unfortunate testing traditions of the past in United States that do not serve the current condition well in the arts. A perspective directed outwards rather than inwards may well serve us better to develop a more productive way forward.

REFERENCES


26. ASSESSING ART LEARNING IN CHANGING CONTEXTS


