

Boughton, D (In Press) Assessment of Performance in the Visual Arts: What, How, and Why. In Karpati, A., Gaul, E. (Eds.) : *From Child Art to Visual Culture of Youth - New Models and Tools for Assessment of Learning and Creation in Art Education*. Intellect Press, Bristol, UK.

ASSESSMENT OF PERFORMANCE IN THE VISUAL ARTS: WHAT, HOW AND WHY?

Douglas G. Boughton, Professor of Art and Education, Northern Illinois University

Assessing the creative outcomes of student learning has always been a dilemma for educators and will continue to be so for as long as we continue to value imagination as one of the most important underlying virtues of engagement in the arts. I have had the good fortune over the past forty five years to spend time teaching art and art education in Australia, Canada, and the United States and since 1993 I have worked as an examiner for the International Baccalaureate program with five years of that time as chief examiner for the visual arts. I have seen many models of assessment in the countries I visited, some of the most interesting of which, to my knowledge, have never been formally studied and documented. The International Baccalaureate itself boasts a robust assessment model that is highly regarded in more than a hundred countries. This chapter is a documentation of what I have learned over the years about the what, who, and how of assessment in the visual arts together with recommendations for best practice.

Assessment, of course, is a continuous process in arts education. The day to day formative judgments made by teachers to assist students' progress towards their learning goals play a central role in any successful art education program. However summative assessment is the force that gives impetus to the direction of the program and defines what is most important to learn. Formative assessment tends to be a private matter between the student and teacher while summative assessments are often a much more public affair. In those countries where art is valued as a core subject the question of how to establish standards is an ongoing debate that finds its way into the public arena. In the Netherlands, for example, the art exam questions are sometimes discussed in the national newspapers.

In this chapter I will discuss summative rather than formative assessment. Good summative assessment models have the potential to improve the quality of student learning in local, state, national, or international contexts also to gain respect for the discipline. I will use the International Baccalaureate model as a backdrop to the discussion of suggested practices. The International Baccalaureate (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 half the countries of the world. As described in documents accessed through the information page of the IB website 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 an effective summative assessment model hinges on three questions "what should be assessed?", "who is qualified to assess it?", and "how should it be done?" I address each of these questions separately in the following.

Boughton, D (In Press) Assessment of Performance in the Visual Arts: What, How, and Why. In Karpati, A., Gaul, E. (Eds.) : *From Child Art to Visual Culture of Youth - New Models and Tools for Assessment of Learning and Creation in Art Education*. Intellect Press, Bristol, UK.

What should be assessed?

For many years educators have argued about specific concepts and skills that should be taught in the visual arts, and which philosophies are most appropriate to direct the selection of such content. I do not wish to enter that debate except to focus on the most problematic assessment issue in the arts, which is how to determine the value of students' creative outcomes.

The assumption that student autonomy is not only important but central to their art making is the assumption that has underpinned art education since the creative free expression movement of the nineteen forties. Student experience in school art programs is thought to develop the capacity for independent thought and the ability to express ideas in visual form. Individual creative expression has long been valued in education systems in most parts of the world. Recent educational reforms, particularly in the United States, have placed significant emphasis upon testing as a way to improve standards in school subjects across the board. An unfortunate by-product for the arts from these reforms in the United States 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, in my view, are the epitome of inappropriate assessment of art learning. These tests require homogenous outcomes reflecting a single set of agreed standards thought to be 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.

While the accountability problems manifested in the United States have not emerged so noticeably in other parts of the world the central assessment question remains equally pertinent. How does one assess students' creative behavior? Creativity is re-emerging as one of the key goals of art learning in the twenty first century (Steers, 2009, Freedman, 2010). Determining the level of a learner's technical skills or knowledge of cultural and historical content is a relatively straightforward task that can be satisfied adequately with traditional assessment methods such as tests, projects, or technical tasks. As we have seen in recent years in the United States the intense focus on multiple choice testing has stripped education to the bare bones of literacy and numeracy and those bones are revealed in schools by multiple choice tests for the most part. Creative behavior cannot easily be measured on a state or national scale so it left to dedicated teachers of art to struggle with the problem alone. In other countries, such as the UK creativity has already been repositioned in the spotlight of educational ambition for the visual arts as is demonstrated by studies initiated at the level of national government (Roberts, 2006; SEED 2006, Steers, 2009).

New ways of thinking about creativity have suggested a pressing need to reconceptualize the art curriculum and the methods used to both assess and promote creative thinking. One of the most powerful, and possibly the most neglected strategies for the development of creativity is the apposite use of assessment. No matter how imaginative or robust new creativity based curriculums might be none will be sufficiently robust to survive inappropriate assessment. By inappropriate I mean the kind of quantitative tests national legislators in the United States currently favor.

As I mentioned previously the current educational policy in the United States following from the *No Child Left Behind* (NCLB) legislation and the conceptions of assessment that followed from this has had devastating impact upon the arts and education generally. NCLB has focused attention upon

Boughton, D (In Press) Assessment of Performance in the Visual Arts: What, How, and Why. In Karpati, A., Gaul, E. (Eds.) : *From Child Art to Visual Culture of Youth - New Models and Tools for Assessment of Learning and Creation in Art Education*. Intellect Press, Bristol, UK.

the so called basics of reading and math diverting attention from social studies, the arts and almost all areas of curriculum that have anything to do with the analysis of human values and cultural issues. NCLB values objective knowledge and answers that are right or wrong. Many schools now take precious time to train students in test-taking strategies. To do this time and resources are stolen from the arts and injected into reading math and science. Even worse, the arts have been cut from the curriculum altogether in some districts.

Creativity is a victim of this pernicious legislation. National testing practices and punishments associated with poor performance for schools unfortunate enough to be under-resourced or populated by large cohorts of special needs students allow no place for creative thinking. Test results and the achievement of national and state specified standards are the only blips on the educational radar in the United States.

In my home state, Illinois I have heard multiple stories from my graduate students about educational administrators in their schools who call upon art teachers to demonstrate the success of their art teaching endeavors with test results showing that students have learned something about art. They expect art teachers to use multiple choice tests to give students practice in filling the bubbles of multiple choice tests. Good art teachers resist such calls, but with sometimes devastating outcomes to their programs.

Creativity Research

Early research in the field of creativity assumed that it was indeed a measurable attribute. Much early research was focused on giftedness and the task of identifying creative individuals. Of particular interest was the development of techniques to identify the personality characteristics and dispositions of creative individuals.

The work of Getzels and Jackson (1971) in the early nineteen seventies revealed some fascinating distinctions between the dispositions and performance of high intelligence students compared to high IQ students. In one study two large groups of adolescent subjects were tested and identified as High IQ or High Creative individuals.

The two groups were compared in a variety of ways to determine differences between them. The researchers found that high IQ students valued personal qualities likely to prepare them for adult success, The highly creative group preferred the opposite. Where the high IQ group favored those qualities they believed the teacher liked, the highly creative group preferred those having no relationship with what they believed would contribute to adult success, and appeared to deliberately select those personal qualities they thought were directly opposite those that their teachers favored (Getzels and Jackson p.127). Creative students it seems are rebellious, uncooperative and nonconforming.

This study is typical of most of the research that was undertaken in the sixties and seventies in the search for answers to the question what are the characteristics of creative individuals? We now have a good set of understandings that help us identify these individuals by virtue of their behavior. But this knowledge does not help us much with pedagogy. If we know creative individuals are nonconforming, and rebellious does it follow that we should try to change the personalities of our students in the hope that they will become creative? Do we direct them to reject authority and seek

Boughton, D (In Press) Assessment of Performance in the Visual Arts: What, How, and Why. In Karpati, A., Gaul, E. (Eds.) : *From Child Art to Visual Culture of Youth - New Models and Tools for Assessment of Learning and Creation in Art Education*. Intellect Press, Bristol, UK.

goals that will not suit them well for adult life? Of course we don't, so what we have to do is look at the problem from a different perspective and this is to pay attention to the environmental conditions that promote creative behavior.

For most of human history ordinary people and researchers alike seem to have attributed creative action to personal attributes rather than the context that promotes creative behavior. Environmental factors contributing to creativity have been largely ignored (Kasof, 1995). To use the distinction specified by Kasof creative behavior has been attributed to *dispositional* rather than *situational* causes. "The result has been a highly skewed research literature in which creativity is studied primarily by personality and cognitive psychologists searching for characteristics of 'creative people' and paying comparatively little attention to external influences on creativity" (Kasof, 1995).

The creative individuals in an art class are not the students that provide the art teacher with their greatest pedagogical challenges. While it is interesting, and perhaps useful, to know how one might identify a creative individual through personality traits, the major concern for the arts teacher is what to do with those individuals who are *not* inherently creative. These students constitute a far larger number in any given class than creative individuals. Is there anything that the teacher can do to promote creative behavior in non-creative individuals?

The most powerful support for educators to examine context rather than disposition comes from the field of social psychology and the work of researchers such as Teresa Amabile (1982), Mihaly Csikszentmihalyi (1996), and Dean Keith Simonton (1979). These researchers have addressed the most pervasive misunderstanding about creativity. That is the subjective reception of a creative product. For something to be regarded as creative it must satisfy two basic criteria. First it must be original, rare, or novel in some way. Second, it must be valued by individuals in the context in which it appears. In other words it must be perceived as approved, accepted, appropriate, or "good." (Kasof, 1995).

By this definition, creativity is not purely objective and is not a fixed attribute of the creative object that holds true irrespective of its time and place. Whether or not an artistic product is creative in part requires a subjective judgment that must be conferred on the original product (Kaslof, 1995; Amabile, 1982; Csikszentmihalyi, 1988 , 1990; Csikszentmihalyi & Robinson, 1986; Gardner, 1993; Weisberg, 1986). As such the determination of creative artistic production becomes an issue of *judgment* rather than *measurement*. It is an *assessment* issue that has a profound effect upon the way art educators need to think about the development of curriculum and the assessment protocols employed for determining student learning.

The fallacy of assuming that creativity is an objective and measurable outcome of learning has significant curriculum implications. Some art curriculums ignore the notion of creativity entirely because of misconceptions about its nature. For example, the current set of state goals in the state of Illinois (USA) does not mention the word creativity because it is difficult to measure through testing.

So what can we draw from the research that is helpful for the art teacher working with the ordinary population of less creative individuals? First we can dispense with the idea that creativity is contingent upon disposition and is therefore dichotomous, i.e. that one is either creative or not.

Boughton, D (In Press) Assessment of Performance in the Visual Arts: What, How, and Why. In Karpati, A., Gaul, E. (Eds.) : *From Child Art to Visual Culture of Youth - New Models and Tools for Assessment of Learning and Creation in Art Education*. Intellect Press, Bristol, UK.

Instead we need to focus upon the classroom conditions that facilitate creative behavior. Second we can set up interrelated curriculum and assessment strategies that promote rather than inhibit creative outcomes, and facilitate judgment processes to determine creative outcomes in a social context.

So what are the conditions that can help to improve creative behavior? Csikszentmihalyi (1995) has offered many positive suggestions for enhancing personal creativity. He interviewed nearly one hundred creative people to gain understanding about the creative process and his recommendations have useful application in the art classroom. They include developing curiosity and interest, cultivating "flow" in everyday life, and ways of thinking creatively. Each of these carries important implications for the creation of conditions that will promote creativity and more importantly defines the nature of assessment that must be used to preserve the integrity of those conditions. I will discuss each of these.

Curiosity and Interest

Csikszentmihalyi says the first step toward a more creative life is the cultivation of curiosity and interest. This seems to be an obvious a suggestion but it is often one that is overlooked. How often have we seen art classes in which students are struggling with media drawing uninteresting or random objects, or simply creating value scales and color wheels for the sole purpose of learning technical processes?

I am not suggesting technical skills should not be taught, or that artists should not be studied. What is of the utmost consequence here, and what is so often overlooked, is the importance of recognizing and engaging the interests students bring to the classroom, and from those leading to new discoveries about technique and artists. Students have a considerable advantage over adults in that their curiosity is easily engaged by many things they encounter in their everyday lives. If invited to bring their interests to the classroom students will willingly oblige. But if art practice is undertaken in the absence of student interest creative production is unlikely to manifest. Interest is an essential prerequisite for creative endeavor.

Journaling and diary notes make experiences more concrete and enduring, and greatly assist students to get in touch with their interests. The point of recording one's experience and surprises is to preserve ideas to make them less fleeting, and after time to look back in order to observe emerging patterns of interest.

Thinking Creatively

Csikszentmihalyi's work also suggests creative thinking is characterized by three fundamental activities, *i) In-depth investigation, ii) Problem finding and iii) risk taking in the search for solutions.*

In-depth pursuit of ideas related to a particular theme is a well documented hallmark of creative behavior. Themes develop from interests and provide unique lenses to view the world thus enhancing curiosity and providing opportunity to develop novel outcomes. Investigating a theme requires work, so there is no point investing energy in a pursuit where there is no interest or passion for discovery. For this reason some people need to explore a variety of thematic investigations before settling on something to pursue in depth.

Boughton, D (In Press) Assessment of Performance in the Visual Arts: What, How, and Why. In Karpati, A., Gaul, E. (Eds.) : *From Child Art to Visual Culture of Youth - New Models and Tools for Assessment of Learning and Creation in Art Education*. Intellect Press, Bristol, UK.

Finding solutions to problems requires divergent thinking and is another key way to engender creative behavior. This is not so much a function of creative disposition as it is a habit of mind. Such habits of thinking can be learned but this requires an individual to consciously seek alternative solutions to a single problem, to experiment, to play, and to take risks. As Elliot Eisner (2005) says producing novelty means one should work at the edge of incompetence. This is risky when you don't quite know what it is you are trying to do but without a supportive and trusting classroom environment risk taking is not likely to occur.

Engendering "Creative Flow"

Csikszentmihalyi is well known for the idea of *flow* which he suggests is the importance of developing habits of engagement with ideas that become self sustaining (Csikszentmihalyi, 1995 p. 349). When one finds an intellectual task that is engaging it is important to be able to pursue it with enthusiasm and sustained interest.

The average school is a very poor place in which to develop creative flow. The structure of a normal school day conspires against the development of any sustained pursuit of creative activity or other kind of intellectual engagement for that matter. In most schools lessons are divided into short time periods of 40 to 80 minutes during which it is scarcely possible for students to collect their materials, let alone their thoughts, in order to generate an idea and begin work. No sooner do they get started the bell sounds and they have to return their materials and move from one classroom to another and repeat the same process over again with a different subject matter. In the course of the day most students start and stop their lessons between five and eight times. Imagine the frustration when one discovers something of interest he or she may wish to pursue only to have to shut down and start something else.

Compared to the intellectual staccato students experience in school, opportunities for out of school visual and intellectual stimulation represent a veritable landscape of treasures. When students can experience the abundance of imagery offered through television, video games, movies, billboards, magazines, the Internet, concerts, exhibitions, community events, and even their phone, it is no wonder they lose interest in school.

There are no easy answers to this problem given the structural limitations of school administration. However, I have seen some hope in the work of gifted teachers who are able to construct the art learning experience as an integral part of the students' life at school. Once interest is engaged in the classroom these teachers encourage students to return during free periods, recesses, lunchtime, and even after the official school day ends. In these classrooms students experience exciting engagement with ideas because their teachers have set up appropriate physical and intellectual conditions, an atmosphere of trust, and the freedom to work in supportive classroom spaces beyond the normal classroom hours.

The Role of Assessment in Fostering Creative Behavior

If we know the situational conditions likely to promote creative behavior then it makes sense to develop assessment strategies that enhance those conditions rather than negate them. We know creative behavior is more likely to occur if curiosity is fostered, if students are encouraged to

Boughton, D (In Press) Assessment of Performance in the Visual Arts: What, How, and Why. In Karpati, A., Gaul, E. (Eds.) : *From Child Art to Visual Culture of Youth - New Models and Tools for Assessment of Learning and Creation in Art Education*. Intellect Press, Bristol, UK.

pursue interests thematically, if they are prepared to play with ideas and engage in risk taking behavior in the search for solutions to problems, and if physical conditions support the idea of creative "flow" described by Csikszentmihalyi'.

I have long argued in support of the use of portfolios as an assessment tool, (Boughton, 2006; Boughton & Wang, 2005; Boughton, 2004; Boughton & Wang, 2002, Boughton, 1996) because good portfolios do more than provide evidence for assessment. They drive curriculum in such a way that creative engagement is more likely. A good portfolio will demand students to demonstrate their interests and show the ways in which they have integrated classroom learning with their lives. A good portfolio will require in-depth and sustained reflection, and will provide a good opportunity to engage interest through the pursuit of thematic content. For a portfolio to have the best chance of becoming a living record of students' creative thinking less assessment is better than more.

The way to destroy creativity through inappropriate assessment is to structure the art program as a series of directed projects that always receive a grade leaving no possibility for a collection of work to be judged as a record of thinking. If the teacher always chooses the topic, the media, the visual references, the reference sources, the strategy, the style of representation, and the look of the potential outcome where is the opportunity for student interests to be engaged? Why would a student take risks in the search for solutions when he or she knows they will be graded on every project they do? Instead, assessment practices that require thematic study, that do not assess each project, that require evidence of productive risk taking, and demand evidence of sustained independent investigation are more likely to encourage creative output.

Who is an authentic evaluator?

If we accept the essence of the argument about creativity offered by social psychologists we must then agree that determining the quality of a creative product is a matter of judgment rather than measurement. The task for teachers is to determine if the student work is original, rare, or novel in some way and that it is valued by individuals in the context in which it is created. This means that teachers need to recognize the social context in which student work is produced. And, if we do that it means the teacher is not necessarily the final arbiter of quality. It does not make sense to ignore the significance of collective judgment about artistic production.

The argument that the teacher should not be the sole arbiter of quality in judging student work is not an expression of mistrust in teachers. Rather it recognises of the nature of art and the ways in which its quality is determined in social settings. Art in the professional world is judged and valued by many in the art community. Critics, artists, agents, and consumers all play their part in stamping an artist's work as original, valuable, worthy or not. A single critic does not make this decision although some may have more influence than others. In the end it is discourse in the social context that establishes the virtue of the work.

Similarly in the educational context there are many stakeholders who can legitimately contribute to the discourse about the quality of artwork made by students. These include the students themselves, the classroom teacher, the community of art teachers, arts administrators, and professional artists to name some.

Boughton, D (In Press) Assessment of Performance in the Visual Arts: What, How, and Why. In Karpati, A., Gaul, E. (Eds.) : *From Child Art to Visual Culture of Youth - New Models and Tools for Assessment of Learning and Creation in Art Education*. Intellect Press, Bristol, UK.

How is he / she selected and trained? How to overcome cultural differences, expert bias or the gap between academic and school practice?

There are some good models to guide us with this collective judgment process. Assessment by students of their colleagues work, and self-assessment within a community context, both help to address the perennial problem of determining the creative quality of artistic products. Moderation processes are employed by school systems in many countries in the world at the senior school level. These have long track records of effectively addressing the need for community determination of the value of art products, and whether or not they contain evidence of creative thinking. More about this later.

In our schools today the demand to demonstrate accountability is extreme and teachers are pressured to produce grades on a regular basis to satisfy the expectations of administrators and parent groups. In the art class this pressure has had the effect of working directly against the development of strategies to enhance creative behavior in art students.

Assessment against the measure of standards has afflicted math, reading and the sciences more particularly than the arts. However the search for ways to achieve predictable and agreed standards in the arts deflects attention away from the search for creative outcomes and the exercise of imagination in our students' art making efforts. Failure to distinguish between standards and standardization in the practice of assessing art destroys the likelihood that students will experience the curricular conditions necessary to stimulate creative thought. It is time to move back towards a more rational relationship between the creative outcomes we desire and the methods we use to assess it.

How to design and conduct summative evaluation?

I have argued above that creativity should be one of the fundamental outcomes of art learning and that an assessment properly conceived and implemented will promote this outcome. The next question is how do we design a summative assessment that will do this in both school and system contexts? The following will focus particularly on the problem of system-wide assessment in contemporary educational settings. By this I mean district, state, or national contexts. The issues to be considered for this purpose include: establishing standards, validity and reliability of assessment judgments in the visual arts, designing authentic assessment tasks, the virtue of portfolio assessment, choosing between analytic and holistic judgment methods, and, moderation models to assist standard setting.

Establishing Standards

The singular interest of system wide assessments is that the achievement of students has common meaning and the standard is understood and accepted by all. The value of an agreed system of standards is that the performance grade awarded to students has common currency. For example students may transport their documented grades from school to school, or school to university, and there is no debate about the meaning. The creation of standards, however, is a most difficult concept for the arts since the notion of "standard" is often confused with "standardization" and its corollary "homogenization". The simplistic and somewhat naive solution (motivated by fiscal

Boughton, D (In Press) Assessment of Performance in the Visual Arts: What, How, and Why. In Karpati, A., Gaul, E. (Eds.) : *From Child Art to Visual Culture of Youth - New Models and Tools for Assessment of Learning and Creation in Art Education*. Intellect Press, Bristol, UK.

restraint) to the accountability problem for the arts in public school contexts in the United States has been paper and pencil testing (Boughton, 2004). Fortunately this is not so much the case in other parts of the world. Nevertheless, it is still possible that the pressure to define and demonstrate the achievement of standards in school art programs could possibly manifest similarly unpalatable outcomes elsewhere if imaginative solutions to the assessment problem are not identified and implemented.

Validity and reliability of assessment judgments in the visual arts

Efforts to standardize assessment content in the arts often violate content validity. What is content validity and why is it threatened by standardized assessment in the visual arts? "Content Validity is based on the extent to which a measurement reflects the specific intended domain of content" (Carmines & Zeller, 1991, p.20). An assessment instrument or practice can be said to have content validity if it is capable of revealing learning that is central to the content of the discipline. For example the demonstration of technical skill and knowledge of formal qualities is easily judged in student work, but possession of this knowledge and skill is peripheral to the core of artistic endeavor. The qualities valued by instructors as the most important defining characteristics of artistic performance are more likely to be attributes such as the degree of imagination exercised, the quality of ideas generated, capacity of students to demonstrate sustained and critical pursuit of themes, and the ability to identify and solve conceptual and technical problems (MacGregor, 1990). These qualities, however, require much more complex judgments from examiners than assessing technical skill and knowledge of the principles of composition.

Herein lies the problem. Accountability pressures to standardize assessment often deflect attention away from the most important content of art learning replacing it instead with content that is most easily and reliably assessed. Content that is easily assessed, however, does not necessarily represent knowledge that is central to the discipline, even though it may be possible to standardize it and measure it with a high degree of reliability. Testing is not a useful solution to the problem despite the reliability payoff. On the other hand traditional portfolios, as an assessment tool, have long offered the potential to achieve high levels of content validity. The portfolio, once the sole province of studio arts, has now been embraced widely throughout the educational and business communities. Although portfolio assessment had its origins in the visual arts it has, interestingly, been recognized as a useful solution to the shortcomings of paper and pencil testing in other subject areas but is largely ignored by both the state high-stakes assessment programs and by teachers of the visual arts in the U.S.A. (Burton, 1998).

Although portfolios offer high content validity in an assessment context, if inappropriate judgment procedures are employed poor inter-judge reliability can result, particularly if individual instructors' judgments are not challenged. Nevertheless, there are some promising solutions to the reliability problem that I will discuss later.

Designing authentic assessment tasks

Much has been written about the concept of authentic assessment in recent years. Howard Gardner (1996) is one who has promoted the concept following from his thinking about the problems of assessment suggested by his theory of multiple intelligences. Authentic assessment

Boughton, D (In Press) Assessment of Performance in the Visual Arts: What, How, and Why. In Karpati, A., Gaul, E. (Eds.) : *From Child Art to Visual Culture of Youth - New Models and Tools for Assessment of Learning and Creation in Art Education*. Intellect Press, Bristol, UK.

(sometimes called performance based assessment, appropriate assessment, or alternative assessment) simply means that students are asked to perform tasks to demonstrate learning directly related to the nature of the discipline with which they are engaged. For example if we want to know if a student can make a realistic drawing of a house it is appropriate to set them a drawing task and then make a judgment about the accuracy and skill of the actual drawing. Asking students to answer questions about drawing, media, vanishing points, or value scales would not be an authentic assessment in this case.

However, if we want to know if a student can make an imaginative drawing of the house the judgment of the artwork becomes more complicated but is, on the other hand, more directly related to the nature of the discipline of art in which we expect creative outcomes.

The virtue of portfolio assessment

While much has been written about the portfolio, it is an instrument that is frequently misused or misunderstood. The portfolio, as an assessment tool, can be conceived in many ways, but in its broadest terms is a body of work collected over time. The real value of portfolios is that it is a time honored performance measure of student achievement, or, in other words, is an authentic assessment.

For assessment purposes the collection of work is regarded as assessment data that will be judged by an assessor or panel of assessors relative to agreed criteria. The first step in achieving reliability of judgment is the development of relevant criteria that reflect content central to the discipline, and these need to be agreed by the community of scholars who constitute the field.

Definition of criteria is not enough however. Assessment of artistic performance represented by work in the portfolio is not a matter of measurement in the same way that knowledge of content can be quantified by multiple choice tests. Assessment of portfolio data requires value judgments to be made about the learning that has taken place and the quality of the work. The nature of the assessment data is therefore critical, so the work within the portfolio needs to be appropriate to the assessment task.

Defining Characteristics of Good Portfolios: The first obvious feature of a good portfolio, mentioned before, is that the work in the portfolio must have been collected over time ... typically a term, semester, or year. Three other important features that define good portfolios are i) that the content is embedded in classroom instruction but remains open ended, ii) the portfolio entries are student selected, and iii) that students document their thoughts (reflections) about their work in verbal and or visual notations. These three features, if overlooked, reduce the potency of the portfolio as an assessment tool.

I will discuss each of these characteristics briefly. The first is that the content of the portfolio is work derived 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 and Herman 1997 use the term *embedded* to describe the way in which students are expected to independently develop ideas and apply techniques that have their origins in the classroom (or studio).

Boughton, D (In Press) Assessment of Performance in the Visual Arts: What, How, and Why. In Karpati, A., Gaul, E. (Eds.) : *From Child Art to Visual Culture of Youth - New Models and Tools for Assessment of Learning and Creation in Art Education*. Intellect Press, Bristol, UK.

The 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 the studio 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. A good example of this kind of portfolio can be seen in the International Baccalaureate program (Boughton, 2004). 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 the art curriculum in very important ways, and is not simply a repository for all class assignments set throughout the year.

Burton (1998) found in a large scale survey that 52% of all visual arts teachers in US public schools assess their students at the completion of each studio project or written assignment. A portfolio that contains only a collection of assigned work and lacks open-ended content is one where the instructor defines both the content and the outcome of each project. Such practice ultimately defines the complete 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 each other 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.

Good portfolios in contrast may contain some common features of students work, but for the most part will be comprised of work that will be unique to each individual, will represent the particular artistic interests of each student, may be very different in content and depending upon the teacher's background, and may even represent a wide array of media as well. Students will certainly have worked outside the classroom and be encouraged to bring their spontaneous work to the studio/classroom to include in their portfolio.

The second feature of good portfolios identified in the literature is that they contain *student-selected entries* (Stecher and Herman 1997; Castiglione, 1996). While the idea of educational portfolios is prominent in the professional art world 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 re-worked 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 their 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

Boughton, D (In Press) Assessment of Performance in the Visual Arts: What, How, and Why. In Karpati, A., Gaul, E. (Eds.) : *From Child Art to Visual Culture of Youth - New Models and Tools for Assessment of Learning and Creation in Art Education*. Intellect Press, Bristol, UK.

this is possible is determined to some extent by the age and sophistication of the students involved. Less is expected of younger students, while more fully resolved work can be anticipated from senior students. Nevertheless, some choice is possible at all levels of schooling. The International Baccalaureate assessment criteria provide useful guidelines (International Baccalaureate Organization, 2000), Purposeful Exploration (Studio), and Independence of Research (Research Workbook) reflect the capacity of portfolios to effectively reveal these qualities in ways that other assessment instruments cannot.

A third, and most important, feature of good portfolios is the significance of *student critical self-reflection*, which may appear in journals or portfolios in written or taped form (Wolfe, 1988). Interviews are commonly used methods in conjunction with portfolios to determine the degree to which students understand their own growth and development. The International Baccalaureate and Arts PROPEL programs in the United States both use this methodology (Blaikie, 1994). Ross, et. Al. (1993) found, during reflective discussions with students, that teachers tend not to listen carefully to students; appear to drive their own agendas through teacher talk; and that students understand more about their own feeling states and sensibilities than their teachers comprehend. Ross claimed that dialogue, properly conducted, can reveal valuable insights into the process of arts 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.

Choosing between analytic and holistic judgment methods

The old adage that the whole is equal to more than the sum of the parts still bedevils judges. One of the perennial dilemmas of judgment strategies in the visual arts is determining the role and value of *holistic judgments* versus *analytic judgments*. Analytic assessment strategies require the specification of discrete criteria which serve to focus attention on those aspects of the work that are thought to be most important. For example, the International Baccalaureate program, prior to 2000, used the following analytic criteria: Imaginative and Creative Thinking and Expression (30%), Persistence in Research (15%), Technical Skill (15%), Understanding the Characteristics and Function of the Chosen Media (15%), Understanding of the Fundamentals of Design (15%), and Evaluation of Own Growth and Development (10%)

The assumption underpinning analytic judgments is that the sum of all the elements defined by these criteria together equals the whole. In some cases application of the analytic method employs criteria that are weighted to reflect their relative significance. For example, the percentages appearing after the International Baccalaureate criteria above indicate that "Imaginative and Creative Thinking and Expression" is twice as important as each of the other criteria except for Evaluation of Own Growth and Development, which is worth only 10%.

The holistic assessment approach, on the other hand, assumes that a single set of criteria cannot be expected to accommodate adequately all genres of visual art work likely to be presented by students in multiple contexts. It may well be the case that the relative emphasis, or type of criteria, employed for appropriate judgement of studio work in different contexts may need to vary somewhat depending on the kind of attention demanded by the work. Indeed, even within a single cultural tradition, the criteria used for judgement may demand different emphases according to its

Boughton, D (In Press) Assessment of Performance in the Visual Arts: What, How, and Why. In Karpati, A., Gaul, E. (Eds.) : *From Child Art to Visual Culture of Youth - New Models and Tools for Assessment of Learning and Creation in Art Education*. Intellect Press, Bristol, UK.

genre. For example, contemporary work using new technologies and recycled imagery, may raise different issues for judgement than work undertaken within traditional styles using older media.

Those who advocate for the holistic approach hold the belief that it is important to avoid the imposition of specific biases, such as cultural or modernist, by the use of criteria that are weighted to reflect a particular view of art. A single set of weighted criteria may be appropriate to some cases, but not to others. For this reason the criteria provided for holistic assessment of studio work are intended only to direct the attention of examiners to values which are important to consider initially in relation to students' work.

Typically the holistic assessment method proceeds as follows. The examiner will first view the work in relation to each of the program criteria. Then the examiner will view the work again to form an overall impression of its qualities considering the particular genre of the work, its cultural emphasis, and any other important characteristics not taken into account by the stated criteria. Finally, a holistic judgment is made taking into account all considerations relevant to the work.

Diederik Schönau (1996) reported at a Getty sponsored Visual Arts Evaluation Conference in Bosschenhoofd the work done by CITO to examine the Dutch Central Practical Examination (CPE) in terms of the potential for achievement of common standards against national prescriptive criteria. Schönau's report of research showed that holistic judgments tended to produce higher judge agreement than judgments using criterion analysis. My own interviews with five chief art moderators in Australia indicated a preference for holistic judgment over strict adherence to criterion based referencing. Judges admit forming an overall impression, then checking the criterion judgments against that overall impression to confirm the final score.

While considerable work is being done in other fields little research is available in visual arts education about the relative value of holistic versus analytic assessments on complex tasks. Other disciplines provide some parallels, for example Walker (1983) found that problems of objectivity in assessing language speaking skills could not be improved with the use of detailed mark schemes. She claimed that the division of language into separately assessed components was inappropriate in oral tests, and did not necessarily provide greater reliability than holistic judgments. In fact, her study confirmed earlier findings by others that dividing oral performance into separate parameters to increase a marker's ability to assess objectively found little difference between the reliability of holistic as compared to analytical mark schemes. Also, the strategy admitted by markers using analytical schemes was to make some kind of general assessment, then apply it to each scale. "Alternatively, experienced markers frequently decide on the total score ("That's a 25 out of 30") and then distribute it among the separate scales." (Walker 1983, p.44). Walker also found that the degree of agreement between judges' holistic impressions was remarkably high.

When I took office with the International Baccalaureate as Chief Examiner for Art/Design in 1994 I raised the question about the appropriateness of an analytic assessment model for art students who are examined across multiple cultural contexts. It seemed to me that the weighted criteria described above represented a universal definition of art that was anchored in a modernist Eurocentric view of art which was not sufficiently flexible to accommodate the full range of work likely to be produced by students in all parts of the world in which the IBO program was taught.

Boughton, D (In Press) Assessment of Performance in the Visual Arts: What, How, and Why. In Karpati, A., Gaul, E. (Eds.) : *From Child Art to Visual Culture of Youth - New Models and Tools for Assessment of Learning and Creation in Art Education*. Intellect Press, Bristol, UK.

Since reliability of examiners judgment is an important precondition for validity in large scale assessments I proposed a research project to investigate the reliability of the analytic model as opposed to a holistic impression of the quality of studio work. The IB assessment system at the time was structured to ensure examiners could not easily check their final score against their holistic impression in order to preserve the integrity of a criterion based analytic assessment system. Pre-determined weightings for each criterion were scaled to a maximum score of five and then the final weighted judgments were computed out of a total seven making it very difficult for examiners to anticipate the final score. This was a highly specified criterion based judgment system that pre-determined relative significance of criteria and disallowed holistic judgment by examiners. The rationale that underpinned this method was that judge agreement would be increased, and this is a very serious concern for any assessment system administrator. The downside was that the method was not as flexible and responsive to emerging new forms of art and design as it should be to achieve the greatest validity of assessment of the lining.

The research proposal to examine the reliability of information gained from weighted criterion based judgments vis-a-vis holistic judgments was undertaken in 1998 as a joint project with Dr George Pook, Director of Assessment for the IB Diploma Program. The research employed a balanced distribution model in which twelve student exhibitions were selected from students in varied cultural settings. These exhibitions, reproduced in slide sets, were split into groups of six and sent to twelve reliable examiners also selected from different world regions. Examiner subjects were identified as those IBO external examiners with a history of reliable judgments over at least three years. Six of these examiners assessed the first half of the slide sets by the holistic method and the second half of the slide sets by the analytic method. The other six examiners did the reverse assessing the first half of the slide sets by the analytic method and the second half by the holistic method. Thus, every collection of students' work was judged by both methods.

The data were analyzed using an inter-rater reliability coefficient to determine the degree of agreement using both assessment methods. It was found there was no significant statistical difference between the reliability of the analytical method as compared to the holistic methods. This study was an IB in-house research project and was not formally published. However the findings were sufficient to enable the examinations office to confidently move to a holistic model for studio assessment which was implemented for first examination in 2002.

The benefits to validity of the holistic model are significant in our post modern age. While "post-modernism" is still not a cohesive, or well defined notion, it is sufficiently evident in the discourse of academic groups within the broader cultural context, to indicate the emergence of a new intellectual mood which has shown itself in the past decade or so. Some philosophers, such as Foucault have claimed that we are experiencing a reconfiguration of Western thought and action on a scale equivalent to the Renaissance. This revolution has been characterized by the rejection of many ideas and practices that began with the Enlightenment of the Eighteenth Century (Shumway, 1989).

Some elements of post-modern thought are already impacting art education. One of the most obvious is the value placed upon the "eclectic" and "bricolage",--- the habit of using whatever comes to hand --- (MacGregor, 1992). The "new architecture" which emerged during the 1970's uses architectural elements from the past, combining them in eclectic configurations that often

Boughton, D (In Press) Assessment of Performance in the Visual Arts: What, How, and Why. In Karpati, A., Gaul, E. (Eds.) : *From Child Art to Visual Culture of Youth - New Models and Tools for Assessment of Learning and Creation in Art Education*. Intellect Press, Bristol, UK.

verge on the bizarre. Jencks (1991) popularised the term "post-modern" in his discussion of these kinds of architectural images.

The practice of plundering the past for inspiration is now as evident in fashion and popular music as it is in other facets of our culture. In the art world ideas from the past now sit comfortably with those of the present. Similarly ideas from different cultures can be accommodated together in the context of pluralistic societies. Many of the old cultural distinctions within, and between cultures are breaking down. In societies where many different cultural groups are living together the cultural activists express their values through art works in curious combinations of imagery and materials. Distinctions that once had meaning are no longer clear or significant. What is the difference between the "original" electronic computer image and the "copy"? Is the distinction important? What is "authentic" in Aboriginal art.. ochre on bodies?... acrylic on canvas?

Other distinctions between high art and popular art are no longer as significant as they once were. In questioning old dichotomies the post-modern age has produced a world of tension between the old and the new, popular culture and high art, conservation and renewal, Western traditions and other cultural practices, in which one is not supposed to be valued over the other.

Several issues arise for educators who have to judge student art learning works, particularly at the senior school level. What is the relationship between post-modernism and modernism in curriculum structure? The issue of originality is challenged by eclecticism, and technical mastery may be even less significant in relation to the idea carried in both visual and verbal forms.

Should the balance in emphasis of socially critical theoretical analysis and studio practice be reconsidered? Aesthetics within a post-modern paradigm may be less significant than social consciousness as a focus for learning. Should the balance of language and image be reconsidered in the total program? Text and image are frequently combined in post-modern expression. At times the image is incomprehensible without the text.

The traditionalist argument suggests established concepts and processes ought to be learned before new expressions can be properly understood. Post-modernism is the newly emerging form, but is only the newest of many other forms of artistic expression which possess their own unique integrity. This debate is a perennial one.

Given the above questions arising from the influence of Post-Modernism there is a very strong case in support of the flexibility of holistic judgments as a means to accommodate the dynamic nature of artistic expression in schools. The value of criteria in judgments is not under question, but strict adherence to them can lead to restrictive conceptions of the field.

Moderation models to assist standard setting

Moderation is a judgment process undertaken by teachers within the educational 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

Boughton, D (In Press) Assessment of Performance in the Visual Arts: What, How, and Why. In Karpati, A., Gaul, E. (Eds.) : *From Child Art to Visual Culture of Youth - New Models and Tools for Assessment of Learning and Creation in Art Education*. Intellect Press, Bristol, UK.

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 in many countries in the world is particularly important to the international Baccalaureate program given the wide geographic distribution of students who participate in the program. Different IB examiners are employed to visit schools in almost half the countries in the world. Subsequent to the examiner visits to schools photographic and photocopy samples of candidates' work is sent to a central location 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. 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.

The benefit of this process, in addition to ensuring more reliable judgments of the quality of student work, is that examiners and teachers receive feedback about their judgments thus developing a community of agreement about standards. The International Baccalaureate is not the only program to employ moderation procedures. Moderation is used on national scale in 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 (Alstete, 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 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

Boughton, D (In Press) Assessment of Performance in the Visual Arts: What, How, and Why. In Karpati, A., Gaul, E. (Eds.) : *From Child Art to Visual Culture of Youth - New Models and Tools for Assessment of Learning and Creation in Art Education*. Intellect Press, Bristol, UK.

qualities rather than specific models of performance. In other words an excellent painting of a scene depicting poverty in Zimbabwe 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 International Baccalaureate 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 website (IBO.org/gallery).

The IB moderation process is based on an external examiner model. That is a model in which teachers do not provide grades for their students' work and an external, impartial expert visits the school to interview students and award the marks. Following the examiners' work a team of moderators will review samples of the examiners marks and make adjustments were necessary to bring them in line with the agreed benchmarks.

There are many other models of moderation. These include peer agreement models in which a committee of instructors review each others' students work. Schönau (1996) reported the system used in the Netherlands in which various committees of peers traveled to schools to review students work at the end of each year. It was found that there was little difference in reliability of judgment between groups of five colleagues and groups of two in which experienced and inexperienced teachers were paired to review each other's schools.

Some school district systems employ direct supervisory review models in which instructors grade work and experienced supervisors moderate. In the United Kingdom a cascade moderation model (Steers, 1996) is employed in which moderators are trained centrally and then in turn train regionally based moderators to work in remote regions. In New South Wales, Australia, senior school students ship their work to a central location in Sydney to be moderated by large teams of centrally located moderators.

Conclusion

In this chapter I have argued that one of the most important and most difficult outcomes of art programs to assess is creative thinking. I have suggested that assessment properly conceived and implemented can foster creative outcomes. To achieve valid assessment outcomes authentic assessment tasks should be employed and one of the most appropriate ways to do this is to employ portfolios both as a source of data for assessment, but also as a central support for the instructional process.

Portfolios need to be appropriately employed to achieve the best results. Good portfolios systematically collect work over time, extend students beyond the classroom, engage student interests, and require a reflective component. Summative, system wide, assessment of student studio work is probably more appropriately conducted by experienced evaluators who employ holistic judgments, clear assessment criteria illustrated by performance descriptors and visual

Boughton, D (In Press) Assessment of Performance in the Visual Arts: What, How, and Why. In Karpati, A., Gaul, E. (Eds.) : *From Child Art to Visual Culture of Youth - New Models and Tools for Assessment of Learning and Creation in Art Education*. Intellect Press, Bristol, UK.

benchmarks. Finally the process of moderation is an essential contributing process for establishing agreed standards, reliability of judgment, and the promotion of healthy discourse in the community of stakeholders in education programs.

Boughton, D (In Press) Assessment of Performance in the Visual Arts: What, How, and Why. In Karpati, A., Gaul, E. (Eds.) : *From Child Art to Visual Culture of Youth - New Models and Tools for Assessment of Learning and Creation in Art Education*. Intellect Press, Bristol, UK.

References

Alsete, J. (1996). *Benchmarking in higher education: adapting best practices to improve quality*. Washington, D. C.: The George Washington University.

Amabile, T. M. (1982). "Social psychology of creativity: A consensual assessment technique". *Journal of Personality and Social Psychology*, 43, (997).

Askin, W. (1985). *Evaluating the Advanced Placement Portfolio in Studio Art*. Princeton, New Jersey: Advanced Placement and The College Board.

Barak, R. & Kniker, C. (2002). Using benchmarking to inform practice in higher education. *Benchmarking by State Higher Education Boards: New Directions for Higher Education*. 18 (Summer).

Blaikie, F. (1994). Values inherent in qualitative assessment of secondary studio art in North America: Advanced Placement, Arts PROPEL, and International Baccalaureate. *Studies in Art Education*, 35 (4), 237-248.

Boughton, D (2004) Assessing Art Learning in Changing Contexts: High-Stakes Accountability, International Standards and Changing Conceptions of Artistic Development. In Elliot Eisner, Michael Day (Eds) *Handbook of Research and Policy in Arts Education*. New Jersey: Lawrence Erlbaum Associates. 585-604

Boughton, D (2006). *Assessment of Student Learning in the Visual Arts* Keynote Address at the Daxia Forum. Published in Chinese Language. Shanghai, China: East China Normal University Publication (September 19th)

Boughton, D. & Wang, S. (2002). The implement of electronic portfolio in student assessment in art education. *Journal of Aesthetic Education* (National Taiwan Arts Education Institute, No.129, 69-75. (in Chinese Language).

Boughton, D. & Wang, S. (2005). Electronic portfolio and visual arts assessment: Conceptual and technical issues in art education. In S. Wang. *Think outside of the box: Visual literacy in the new century*, Taipei: National Taiwan Arts Education Center (in Chinese Language). 48-57.

Boughton, D. (1996). Evaluating and Assessing Art Education: Issues and Prospects. In Boughton, Doug.; Eisner, Elliot.W.; Ligtvoet, Johan. *Evaluation and Assessment of Visual Arts Education: International Perspectives*. NY: Teachers College Press. 293-310

Boughton, Doug (1994) *Evaluation and Assessment of Visual Arts Education*. Geelong: Deakin University Press.

Boughton, Doug. (1997) Reconsidering Issues of Assessment and Achievement Standards in Art Education. *Studies in Art Education*, Vol 38, No 4, 199-213.

Boughton, D (In Press) Assessment of Performance in the Visual Arts: What, How, and Why. In Karpati, A., Gaul, E. (Eds.) : *From Child Art to Visual Culture of Youth - New Models and Tools for Assessment of Learning and Creation in Art Education*. Intellect Press, Bristol, UK.

Burton, D. (1998). *A Survey of Assessment and Evaluation among U.S. K-12 Teachers of Art*. Meeting of the NAEA Task Force on Demographic Research. NAEA convention, Chicago IL., April 2.

Carmines, E. G. & Zeller, R.A. (1991). *Reliability and validity assessment*. Newbury Park: Sage Publications.

Castiglione, L. (1996). Portfolio Assessment in Art and Education. *Education Policy Review*, 97 (2), 2-9.

Chapman, Laura. (2005) No child left behind in art? *Art Education*, 58 (1) 6-16

Codling, S. (1998). *Benchmarking*. Brookfield, Vermont: Gower.

Csikszentmihalyi, M. (1996) *Creativity: Flow and the Psychology of Discovery and Invention*. NY: Harper Collins Publications.

Csikszentmihalyi, M. (1988). "Society, culture, and person: A systems view of creativity". In R. J. Sternberg (Ed.), *The nature of creativity: Contemporary psychological perspectives* 325-339. Cambridge: Cambridge University Press.

Csikszentmihalyi, M. (1990). "The domain of creativity". In M. A. Runco & R. S. Albert (Eds.), *Theories of creativity* (pp. 190-212). Newbury Park, CA: Sage.

Csikszentmihalyi, M., & Robinson, R. (1986). "Culture, time, and the development of talent". In R. J. Sternberg & J. E. Davidson (Eds.), *Conceptions of giftedness* (pp. 264-284). New York: Cambridge University Press.

Eisner, E. W. (2004) Response to Arthur Efland's and Richard Siegesmund's Reviews of The Arts and the Creation of Mind, *Journal of Aesthetic Education*. 38, 4. pp.96-98

Eisner, E.W. (2005) *Reimagining schools: The selected works of Elliot W. Eisner*. New York: Routledge.

Freedman, K (2010). Rethinking Creativity ; any definition two support contemporary practice. *Art education*, 1:

Gardner, H. (1993). *Creating minds*. New York: Basic Books.

Gardner, H. (1996). *The assessment of student learning in the arts*, In Boughton, Doug.; Eisner, Elliot.W.; Ligtvoet, Johan. *Evaluation and Assessment of Visual Arts Education: International Perspectives*. NY: Teachers College Press 131-155.

Getzels, J., & Jackson, P. (1971) The Highly Intelligent and the Highly Creative Adolescent: A Summary of Some Research Findings, in Rolf E Muuss (Ed,) *Adolescent Behavior and Society: a Book of Readings*. NY: Random House Inc.

Boughton, D (In Press) Assessment of Performance in the Visual Arts: What, How, and Why. In Karpati, A., Gaul, E. (Eds.) : *From Child Art to Visual Culture of Youth - New Models and Tools for Assessment of Learning and Creation in Art Education*. Intellect Press, Bristol, UK.

International Baccalaureate Organization. (2000) *Diploma Program Visual Arts Guide: For First Examinations 2002*. Geneva, Switzerland: The International Baccalaureate Organization.

Jencks, C. (1991). Post-modernism in London. Architectural Design, Academy, London.

Kasof, J. (1995). Explaining Creativity: the Attributional Perspective. *Creativity Research Journal*.8 (4), 311.

MacGregor, R.N. (1990). Reflections on art assessment practices. *Journal of art and design education*.

MacGregor, R.N. (Sept, 1992) Post-Modernism, Art Educators, and Art education. ERIC Digest ED348238. Bloomington, IN: ERIC Clearinghouse.

Roberts, P. (2006) *Nurturing Creativity in Young People: A report to Government to inform future policy*. Department of Culture Media and Sport, UK: (July) <http://www.ltscotland.org.uk/creativity/aboutcreativity/nurturingcreativity.asp>

Ross, M., Radnor, H., Mitchell, S., Birtton, C. (1993). *Assessing Achievement in the Arts*. Buckingham, UK: Open University Press.

Schönau, D (1996) Nationwide assessment of studio work in the visual arts: Actual practice and research in the Netherlands. In Boughton, Doug.; Eisner, Elliot.W.; Ligtvoet, Johan. *Evaluation and Assessment of Visual Arts Education: International Perspectives*. NY: Teachers College Press.156-175.

SEED [Scottish Executive Education Department Promoting Creativity in Education], (2006): *Overview of Key National Policy Developments Across the UK*. An Information Paper by SEED, (September) <http://www.hmie.gov.uk/documents/publication/hmiepcie.html>

Shumway, D. (1989) Michel Foucault. Boston: Twayne. Twayne Masters Series.

Simonton, D. K. (1979). "Multiple discovery and invention: Zeitgeist, genius, or chance?" *Journal of Personality and Social Psychology*, 37, 1603-1616.

Stecher, B., & Herman, J. (1997). Using portfolios for large scale assessment. In Gary D. Phey *Handbook of Classroom Assessment*. London: Academic Press Limited. 491-516.

Steers, J (2009) Creativity: Delusions, Realities, Opportunities and Challenges, *International Journal of Art & Design Education*, 28, (2) 126-138.

Steers, J. (1996) Response to Papers by Gardner and Schönau. In Boughton, Doug.; Eisner, Elliot.W.; Ligtvoet, Johan. *Evaluation and Assessment of Visual Arts Education: International Perspectives*. NY: Teachers College Press, 176-193.

Steers, J. (1988). Art and design in the national curriculum. *Journal of Art and Design Education*. 7 (3), 303-323.

Tucker, S. (1996). *Benchmarking: a Guide for Educators*. Thousand Oaks, California: Corwin Press Inc.

Boughton, D (In Press) Assessment of Performance in the Visual Arts: What, How, and Why. In Karpati, A., Gaul, E. (Eds.) : *From Child Art to Visual Culture of Youth - New Models and Tools for Assessment of Learning and Creation in Art Education*. Intellect Press, Bristol, UK.

Walker, C.R. (1983 Spring) The whole is more than the sum of the parts: Evaluation criteria in oral tests. *British Journal of Language Teaching* 21 (1) 41-44

Weisberg, R. W. (1986). *Creativity: Genius and other myths*. New York: Freeman.

Wolfe, D. (1988). Opening up assessment. *Educational Leadership*, 45(1), 24-29.