Assessment and Visual Arts Education

Elisabeth Soep
Youth Radio & University of California, Berkeley

The relationship between art and assessment is best characterized as awkward, if not overtly hostile. Arts educators tend to bemoan the impact of mandatory evaluations on imaginative practice. To make matters worse, accountability systems and testing conventions are often imported from academic subjects—domains of learning that may more easily “break down” into component questions with clear right answers, which students can conscientiously bubble in on multiple-choice tests. Visual arts projects can be messy, whether they involve hands-on production tasks or responses to work by established artists. It is difficult, and some might argue damaging, to evaluate student performance on these kinds of tasks according to predetermined standards. Matters get even more complicated when judgments about the quality of student work go beyond individual encounters between teachers and their students, when educators are answerable to state-sponsored mandates or external reviews.

This view, which pits art against assessment like two opponents in a boxing ring, is often justified. Nevertheless, much can be learned when we force ourselves to recognize what the work of visual art and the work of assessment have in common. Perhaps the two are not as oppositional as they initially appear.

1. Artworks and assessment works visualize the ineffable. Both visual artists and assessors are often in the business of rendering tangible something that defies easy articulation. Now, one could argue that plenty of tests merely require students to report bits of information—correct dates, names, and solutions, for example. But underlying individual test items is usually a drive to understand something much more elusive, like student learning, or what a young person knows and is able to do. Likewise, artists are not always interested in conveying some profound, ground-breaking message. Yet artworks function in a larger sense as objects that capture and convey complex meanings (even when they call into question the very possibility of “meaning” as an attribute of art). Both assessors and artists operate in worlds where tensions of translation are inevitable, in the always imperfect action of turning thoughts into things. In this section of the Handbook, Persky in particular explores these tensions, in her analysis of efforts by national assessors not to overburden visual arts tasks with verbal directions—a
580  SOEP

problem of translation across symbol systems. Despite or perhaps by virtue of these kinds of
tensions, artworks and assessment works partially articulate some truth, or some lie, that would
otherwise have remained invisible, had no effort taken place to make the ineffable concrete.

2. **Artworks and assessment works tell stories.** One way to look at individual art objects,
collections of related works, or trajectories whereby whole traditions have emerged, dominated,
and then fallen out of favor is to ask what stories these artifacts tell. The same is true for
assessment. This point surfaces especially in one essay in this section by Myford and Sims-
Gunzenhauser, where the authors describe the evolution of two national visual arts evaluation
programs as a “story of contrasts” that reveals how educational priorities and assumptions
have changed from the 1970s to the present. Art objects and histories, like assessment contents,
structures, and rationales, expose who we are, as individuals and societies, at any given moment
in time. Even the most antinarrative artwork and even the most statistics-driven evaluation have
important stories to tell.

3. **Artworks and assessment works provoke controversy.** Related to point number 2 is
this third connection: that art and assessment can tell us things about ourselves that we do
not want to know. This observation seems obvious enough with respect to artworks. Societies
have long tried to silence artists who expose painful realities or defy mainstream sensibilities,
just as we depend on artists to advance avant-garde ideas and sensibilities. This same quality
of provocation can be attributed to assessment, on at least two levels. First, there are the
public outcries and “moral panics” that follow any large-scale report of student failure on
standardized tests; often poor outcomes give rise to heated debates, new policy and curriculum
interventions, and, predictably, ever more energetic testing efforts. Second, the difficult and
provocative nature of assessment operates on a deeper level as well—when tests are exposed to
be biased, or inhibiting to curricular innovation, or abused when applied to malevolent purposes.
Artworks, and assessment works, can do harm or make good; both can yield uncomfortable
moments of recognition when we are forced to contemplate our failings as artists and viewers,
as test takers and test makers, and as participants in institutions that shape both realms of
practice.

4. **Artworks and assessment works simultaneously reflect and create the people who
encounter these cultural products.** The Brechtian idea that artworks are not merely mirrors
held up to a society but hammers that shape it holds true as well for assessments. Artworks
and assessment works are *technologies*, influenced by new materials and sources of knowl-
dgeability, and also generative of these innovations through experimentation—for example,
when affordable editing software became available to the grassroots video community, or with
the advent of digital portfolios as a facet of visual arts assessment. The tools we have at our
disposal not only serve or measure but also help shape what we make and know. Artworks
and assessment works are also *ideologies* based in specific worldviews about the kinds of
efforts worth undertaking and recognizing, as well as those deserving of little notice or loud
resistance. It is possible to call an assessment tradition, like an arts tradition, “modernist,” as
Boughton does in his chapter for this section, with respect to standards-based multiple-choice
testing. And there is growing interest today in what “postmodern” assessment might look like
in contemporary classrooms (Fehr, Fehr, & Keifer-Boyd, 2000). Artists and assessors belong
to, sustain, and ultimately unsettle specific movements and schools of thought and action. Just
as “tests have invented all of us” (Hanson, 2000), so too do artworks transform those who
make, consume, and critique visual culture.

5. **Artworks and assessment works are social practices.** Despite the persistent myth
of the artist as a lone genius working under sequestered conditions, most artists and arts
educators will attest from experience that artworlds are social fields, populated by peers, critics,
mentors, models, and institutions including museums, popular culture industries, and schools.
Even private moments of inspiration operate within minds and bodies shaped by histories of
interaction with others. This same point about the social character of arts experience can be applied to assessment. Although it may seem intuitive to locate assessment projects squarely in the realm of scientific practice, even statistical methods and empirical findings develop out of social processes, which define quality and merit within a given realm of performance (Broadfoot, 2000). In this section of the Handbook, I make this sociocultural approach explicit, through my review of research on self- and peer assessment in visual arts education—processes that can range from formal and high-stakes “peer reviews” to informal and fleeting moments when a child comments on her classmate’s drawing. Viewing artworks and assessment works as social practices draws attention to the complex orchestration of ideas, values, and institutional sanctions that support certain traditions within both fields. This approach also points to the possibility of intervening within these fields, by reminding us that these two realms of social practice are always in a process of reinvention.

Identifying these five features that artworks and assessment works have in common should in no way diminish the tensions that arise when arts educators are called on to evaluate their own pedagogical practices and their students’ efforts and products. These conflicts are very real, and very significant, as the arts find themselves excluded from the vast majority of state-evaluation programs in the United States, even as arts educators are constantly expected to justify their place in the curriculum. This position is both a blessing and a curse. The arts are protected, to some extent, from the incursion of rigid standards and high-stakes pressures, but they are also overlooked in resource allocation and public recognition, thereby potentially reifying a marginalized place for the arts among school subjects.

The first author in this section, Doug Boughton, begins his chapter by exposing this odd position for the visual arts—a field that is rarely required for college acceptance, included in only seven states’ mandatory testing policies, and yet caught up in a climate that demands accountability and public scrutiny for all subject areas that really matter in education. Boughton describes what he sees as a promising model for large-scale arts assessment administered in over 60 countries around the world—the International Baccalaureate (IB) program. In the IB program, trained external examiners assess portfolios documenting students’ research and studio practices using predetermined holistic categories. Boughton argues that the portfolio is a far more appropriate measure of arts learning than paper-and-pencil tests, which “dismember” artistic performance and thereby stand to do more harm than good. The chapter holds relevance beyond the IB program in particular, as Boughton provides an overview of portfolio assessment in the arts, identifying the features associated with learning portfolios as distinct from professional portfolios. To function as an assessment tool within visual arts education, the portfolio should be more than a “repository” for homework assignments; Boughton says students should be the ones creating portfolio archives, including works in progress, sketches, and their own critical self-reflections. Two additional dimensions of the IB program also have broad relevance to debates in visual arts education: (1) A process called grade moderation, whereby multiple judges assess student work, promoting a climate of debate and a community of deliberation; and (2) an approach for developing assessment benchmarks based on exemplar student work embodying certain qualities, and not based on models put forth to be copied. Overall, Boughton makes a case for portfolio assessment in the arts without downplaying the tensions likely to arise when external measures are introduced within a U.S. arts educational climate that tends to privilege informal teacher judgments of student work.

The two subsequent chapters in this section of the Handbook also address large-scale arts assessments. Hilary Persky reports on the 1997 National Assessment of Educational Progress (NAEP) visual arts assessment. She sees NAEP as a model for capturing individual meaning-making and artistic thinking, assuming that the arts are integral to education “for all” and
not only for the few who self-select or happen to attend schools with strong arts programs. Unlike the IB program, NAEP is a one-day, “drop-down,” timed assessment for which students have no specific preparation. The visual arts section is just one of several academic and arts domains included in NAEP, which is administered to representative random samples of students across the United States, including those with absolutely no arts training. Needless to say, these conditions create unique challenges for designing a valid, reliable, and feasible evaluation constrained significantly by concerns about time limits, task sophistication, and labor and material costs. NAEP only works if the evaluators can get in and out of schools in a matter of hours, having never before met the students and knowing nothing about the specific contexts of the classrooms, schools, or communities where they are administering the test. It is a daunting undertaking indeed, says Persky, to devise tasks that allow students to create two- and three-dimensional objects, as well as to respond to the work of established artists, using various materials, all in the space of a few hours. Student work generated for NAEP is shipped off to external, trained judges for scoring, adding yet another significant challenge. Persky’s chapter presents a case where assessors aim to recognize varied, imaginative, and experimental student responses despite fixed scoring criteria and strict format constraints. An additional contribution of the chapter is Persky’s inclusion of several actual examples of student work—written and visual—produced in response to NAEP tasks, offering readers the artifacts themselves to consider alongside the scores these pieces actually earned, revealing the strengths and weaknesses of one of the most ambitious national visual arts assessments this country knows.

Persky’s chapter provides a nice lead-in to the next one by Carol Myford and Alice Sims-Gunzenhauser, which chronicles the evolution of both NAEP and another national arts assessment, the Advanced Placement (AP) studio art program, over 3 decades. With respect to NAEP, the authors essentially compare the 1997 program with prior administrations of the assessment in 1975 and 1978. The 1970s favored breadth of coverage and behavioral objectives, whereas the 1997 version emphasized depth of knowledge and skill, range in tasks, and specific process and content standards. The most recent administration of the NAEP also showed increased sophistication in test design and scoring technologies and had a much larger evident impact on the research community. The AP assessment, unlike NAEP, has been given continuously since 1972, based on cumulative student portfolios rather than drop-down timed testing events. Over the years, the structure for the AP has shifted, but the evaluation of student portfolios has consistently depended on highly trained “readers,” themselves high school- and college-level art teachers, who collaboratively review submissions based on holistic measures. Beyond analyzing the specific cases included in this chapter, Myford and Sims-Gunzenhauser address the more general tensions surrounding any assessment effort that extends over a significant period. The assessment needs to evolve and transform with the times. However, changes to the assessment can prevent comparison across groups of students who have taken different versions of the test (or experienced different approaches to the portfolio, or tackled different kinds of tasks). Herein lies the profound assessment challenge of balancing benefits for student learning and updated practice with social science principles that dictate valid, reliable, and significant comparisons among young people learning in diverse contexts. Myford and Sims-Gunzenhauser, in their discussion of the AP exam for studio art, also provide a glimpse into the kinds of deliberations that take place when panels of readers jointly assess work submitted within student portfolios.

In the final chapter of this section, I focus on the process of joint assessment among peers, as well as on individual self-assessment, by students themselves. The chapter begins with a discussion of relevant theories pertaining to self- and peer assessment in the visual arts, proposing a sociocultural view of the mind and outlining the notion of “artworlds” that assign value to specific arts practices while rendering others less visible or desirable. Organizing the
chapter roughly along a developmental trajectory, I cast a wide net, reviewing research drawn from anthropology, cultural studies, education, and psychology that sheds light on how young artists formulate judgments of their own projects and those of their peers. As a complement to the other essays, which center on standardized national assessment programs, this chapter tends to zero in on small-scale, moment-to-moment assessment encounters among learners, and between students and teachers, which may or may not take place within formal evaluation events or accountability programs. This view of assessment as an unfolding process takes readers inside classrooms as well as to less recognized arts learning environments beyond school walls, expanding what counts as arts education in an era of increasingly pervasive and complex visual cultures.

Overall, the chapters contained in this section do not shrink from acknowledging the difficulties inherent in assessing visual arts learning. Perhaps one conclusion readers might draw is that there is, in fact, much to moan about when educators are asked to evaluate what their students know and are able to do in the arts. An alternative message to glean from these authors is a view of the unique challenges faced by arts educators not as burdens, but as opportunities. Every field of learning—including the arts and math and science and English—features relatively measurable skills as well as sophisticated performances that defy easy classification, especially when students make efforts that are truly experimental and unexpected. When assessments capture this full range of student abilities (as Myford and Sims-Gunzenhauser argue in their chapter), the dreaded act of “teaching to the test” actually results in “students working with commitment and passion at visual issues that stretch their knowledge, strengthen their conceptual abilities, and, of course, help them develop the framework of technical skills that are necessary to realize their ideas.” Arts educators have no choice but to aim for these ambitions, simply because of the nature of our field. That surely makes the work harder, but it also puts the arts in a position of leadership with respect to innovative education, and assessment, across disciplines.

REFERENCES


