Handbook of Research and Policy in Art Education

Elliot W. Eisner, Michael D. Day

The NAEP Arts Assessment: Pushing the Boundaries of Large-Scale Performance Assessment

Publication details
Hilary Persky
Published online on: 23 Feb 2004


PLEASE SCROLL DOWN FOR DOCUMENT

Full terms and conditions of use: https://www.routledgehandbooks.com/legal-notices/terms

This Document PDF may be used for research, teaching and private study purposes. Any substantial or systematic reproductions, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The publisher shall not be liable for an loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.
INTRODUCTION

In 1997, the national assessment of students’ visual arts knowledge and skills provoked both excitement and trepidation among many arts educators and educational policymakers. These sentiments were expressed throughout the process of assessment development, administration, and scoring by members of the National Assessment of Educational Progress (NAEP) visual arts development committee, many of whom had been part of the creation of the arts voluntary national standards and the NAEP arts education framework. Excitement and pride at the prospect of a national arts assessment combined with valid concerns about how fairly a standardized assessment could elicit and measure art making. As stated in the 1997 Arts Education Framework that served as a blueprint for the NAEP arts assessment, “many arts educators worry that an assessment of the arts will artificially quantify the aspects of the arts that seem unquantifiable—inspiration, imagination, and creativity” National Assessment Governing Board [NAGB], (1994). (Audiences also frequently expressed this concern when my colleagues and I spoke to various state arts assessment organizations about the national assessment initiative.)

In spite of these concerns, the 1997 NAEP arts assessment was highly productive. A determination on the part of its creators both to push and to respect the limits intrinsic to national large-scale assessment made for an arts assessment that meaningfully informed the general public about the state of arts education in the schools. Further, many state arts assessment initiatives, including those in Maine, California, New Jersey, New York, and Maryland, have drawn and continue to draw on the national arts assessment model for item development, administration, and scoring. All five states requested presentations and workshops about the NAEP arts assessment, and information about how NAEP arts items were created and utilized was very positively received. Visual arts tasks that combine responding to works of art with creating tasks have been adapted from the national assessment model for use New Jersey, Maryland, and New York, that the author is aware of.
More broadly, the 1997 visual arts assessment provides a solid model for capturing the individual, expressive nature of making meaning, in ways that simulate good classroom practice, and still yield comparable, meaningful results from students with a wide range of exposures to arts education. Still, some would argue that NAEP was only partially successful in meeting these challenges (Beatty, 1999; Eisner, 1997). In presenting several strategies that designers and advisors adopted for the NAEP arts, this chapter will reflect on how far the 1997 effort took the arts assessment field. The final section of this chapter argues that although success at a national level is intrinsically limited, NAEP can offer informative and useful models for new approaches to large-scale arts assessment.

BACKGROUND TO THE NAEP 1997 ARTS ASSESSMENT

Since 1969, the National Assessment of Educational Progress (NAEP), also known as “the Nation’s Report Card,” has been the only nationally representative and continuing assessment of what America’s students know and can do in various subject areas, including reading, mathematics, science, the arts, writing, U.S. history, civics, and geography. Under the current structure, the Commissioner of Education Statistics, who heads the National Center for Education Statistics (NCES) in the U.S. Department of Education, is responsible by law for carrying out NAEP projects. The National Assessment Governing Board, appointed by the Secretary of Education, but independent of the department, governs the program. NAEP does not provide scores for individual students or schools. Instead, the test offers results regarding subject-matter achievement, instructional experiences, and school environment for selected national populations of students (e.g., fourth-graders) and subgroups of those populations (e.g., female students, Hispanic students).

The NAEP 1997 arts assessment was the result of a multi-year process, described briefly in this volume (Myford & Sims-Gunzenhauser, 2002). Myford and Sims-Gunzenhauser discuss the collaboration among arts educators, artists, policymakers, and members of the public in the creation of the voluntary National Standards for Arts Education and The 1997 NAEP Arts Education Assessment Framework. This collaboration resulted in two documents that share a set of ideas, although the standards describe what ought to be taught in the arts, and the framework describes what and how to assess. Both documents assume that an arts education is not just for the talented or the privileged, but is instead an integral part of education. As stated in the framework:

Throughout their lives, [children] will draw from artistic experience and knowledge as a means of understanding what happens both inside and outside their own skin, just as they use mathematical, scientific, historical, and other frameworks for understanding. The expectation is not that they will become talented artists. What is expected is that they will have experienced enough of the discipline, the challenge, the joy of creating in different art forms to intimately understand the human significance of dance, music, theatre, and the visual arts. (NAGB, 1994, p. 1)

The notion that an arts education is not meant primarily to create talented artists should not be taken to mean that arts knowledge and skills are simple and easily attainable. Like other disciplines, dance, music, theater, and visual arts can be extremely challenging, and any assessment of arts knowledge and skills ought to elicit high levels of performance, not just minimum competency. Further, the standards and framework express the following principle: If a sequential and rigorous arts education is a crucial part of the curriculum, and such an education must emphasize creating, performing, studying and analyzing works of art, then any arts assessment ought to include chances for students to analyze, critique, and formulate value judgments about works of art, as well as to create and perform works of art.
Given these ambitions for a rigorous arts assessment, the tension between the depth and continuity of what can happen in an arts classroom and what can be managed in a large-scale arts assessment is a central concern of the arts assessment framework. In a classroom setting, teachers can evaluate students’ arts knowledge and skills on the basis of prolonged observation. Students can ask questions and discuss artworks, their ideas, and artistic choices with peers and teachers; explore and experiment with different strategies for creating; and work on their projects over time. This is not the case in a drop-in timed assessment, for which students have had no specific preparation and that must be completed within a short period on a single day.

The task for those who created the NAEP visual arts assessment was therefore to encourage students to reflect about works of art, communicate ideas and feelings about works of art, and to take imaginative risks to solve artistic problems. At the same time, the assessment needed to be practical to administer and to yield responses that could be compared and scored on a large scale. This last was especially difficult because the students who were assessed had many different arts backgrounds. Teaching approaches to the arts vary widely. Further, because the arts are often not part of standard school curricula, some students had no arts background, whereas others had quite substantial arts experience, gained either in or outside of school. (The challenges of making assessment tasks accessible to a very wide range of students and of locating student samples that can meaningfully perform on an arts assessment are challenges that are somewhat peculiar to NAEP, as NAEP is mandated to assess representative national samples of students.)

To ensure that assessment tasks were as valid and reliable as possible, a team of arts experts was assembled to work alongside measurement specialists at Educational Testing Service (ETS). (ETS currently holds the NAEP development, analysis, and reporting contract.) This committee was composed of arts classroom teachers, arts curriculum experts, artists, and arts policymakers, many of whom had worked on the voluntary standards and the NAEP arts framework creation. Committee members brought their expertise to creating an arts performance assessment that would be as “authentic” and as valid as possible. This meant that all assessment tasks were created, reviewed, discussed, and refined by arts education experts who sought to adhere as closely as possible to the visions of the arts standards and framework; the goal was to develop exercises and scoring criteria that would measure important arts knowledge and skills in grade-appropriate ways. This included selecting and approving sets of exemplar students’ works to be used to train raters to score assessment results. Some of the challenges confronted by ETS staff and the committee and the pros and cons of the strategies adopted to meet those challenges are explored in the following section.

Issues and Strategies Used in Assessment Development and Administration

This section will explore how visual arts tasks were developed and organized to accommodate administration concerns while being responsive to the arts framework. NAEP item development must always take place with an eye toward what will actually happen in schools. The most valid assessment task will fail if it cannot be administered, either because the task materials and space requirements are too complex or because the task itself is too complex to be understood and completed by students within a limited time frame. NAEP is a voluntary national assessment and must compete with many state-level assessments and standard curricula for student’s time. Securing school participation in NAEP is often dependent on getting in and out of schools rapidly and as efficiently as possible; NAEP administrations thus cannot consume too much of school time or place undue burdens on schools for space or complex movements of students from one space to another.
To address the problem of balancing assessment depth with the time and space that could reasonably be asked for from schools, NAEP created separate assessments for the four arts subjects. This meant that any individual student participating in the assessment was assessed in only one arts subject, that is, dance, music, theater, or visual arts. Assessing each student on an individual basis in each subject area would have been prohibitively expensive and would have created massive time burdens on schools. Individual students usually spent an average of 90 minutes engaged in tasks for one arts subject, allowing a relatively in-depth assessment of the students’ abilities in that subject. Separate assessments for each arts area were a natural first step to gain assessment depth (although it did preclude another framework goal of exploring relationships among the arts areas).

The arts framework also sought assessment authenticity. In visual arts, for example, students were to be offered a range of media and engage in a range of art-making processes including design. Further, arts processes were to be understood as inseparable: “The NAEP framework and the National Standards have framed a vision of arts education that integrates the aesthetic, social cultural, and historical contexts of the arts with the knowledge and skills necessary to participate in the arts. Skills will not be considered as separable...” (NAGB, 1994, p. 3). “The assessment will consist largely of multiple, related, exercise organized around an activity” (NAGB, 1994, p. 4). And finally, “Exercises should include a mix and a balance of creating exercises and responding exercises that engage a wide variety of knowledge and skills in studio production, art criticism, art history, and aesthetics” (NAGB, 1994, p. 7).

As will become evident from the following discussion, combining these processes, knowledge, and skills to give students enough time to create and respond meaningfully posed school burden challenges both in terms of time and space. Typically, students taking the NAEP in any subject area respond to two “blocks” (sets) of questions to be completed within a set time period (usually 50 minutes). Students complete anywhere from 15 to 24 assessment exercises, a combination of enough multiple choice and open-ended items to allow NAEP scaling and analysis. To enable efficient administration while being responsive to the framework, ETS staff and visual arts committee members decided to combine responding items requiring written responses to artworks with two-dimensional creating tasks, and then developed entirely separate tasks for three-dimensional creating. Both kinds of tasks were embedded in “focused” blocks concerned with a particular problem, issue, or genre. The logic and success of this design for the NAEP visual arts are discussed in the following section.

Challenges Associated With Administering Two- and Three-Dimensional Visual Arts Tasks

In developing assessment items for visual arts, pretesting revealed that a single block designed to measure students’ responding abilities alone took anywhere from 40 to 50 minutes to complete, given the time students needed for absorbing visual stimuli and answering thoughtful questions in writing. If students then had to be given time for creating activities, the time for assessing any one student could easily extend to well over 2 hours. Two hours of assessment time would be compounded by time for setup, materials distribution, the gathering of materials, and cleanup, thus increasing assessment time for a set of visual arts tasks to a half-day or more.

Administering three-dimensional creating tasks posed a set of additional and unique challenges. Although two-dimensional works could be collected and shipped for eventual scoring, three-dimensional artworks needed to be photographed. It would have been impossible, given the realities of time, equipment, staff, and thousands of fragile artworks, to carefully wrap each three-dimensional work to ship it for scoring. Hence, administrators were taught during administrator training how to photograph each student’s three-dimensional work from several different angles for scoring. Although this was surprisingly successful, it was certainly limiting;
it is always preferable to be able to look at three-dimensional works from every angle. There was some discussion of having experts score on site. However, organizing experts to be on site at the right time, given a nationwide assessment in which it is better not to use teachers of students at any given school to score, was problematic. NAEP prefers to train groups of expert, independent raters to agree on complex scoring criteria to demonstrate the validity of student responses.

Even apart from the photographing of student artworks, the complications involved in handing out and accounting for the complex materials necessary for three-dimensional tasks and the space required by students to create effectively with such materials made the tasks both very expensive and time consuming to administer. Two-dimensional tasks were difficult enough to administer; administrators had to be taught how to account for and distribute carefully packaged arts materials. Each kind of media and tool had to be separately bagged or bagged in sets, enabling administrators to easily check that a given school had received the correct shipment, and that each student received the appropriate number of and kind of media and tool. As an example, shown in the following is the materials list for a collage task that combines responding and creating activities. The list demonstrates the challenges of assessment administration caused by the justifiable efforts of assessment developers to make tasks as valid and as authentic as possible by supplying suitable ranges of materials for student creating. Students, after observing a Romare Bearden collage and answering questions about it, created their own collages, and each student required the following materials:

- **Test booklet**
- **Pencil kit** (one 2B pencil, one pink pearl eraser, one handheld sharpener)
- **One portfolio containing:**
  - Glue stick
  - Black ball-point pen
  - 12 in. × 18 in. 80 lb. drawing paper
  - 4 in. × 5 in. postcard of Bearden collage
  - Set of 12 Cray-pas
  - Set of 8 fine-tip watercolor markers
- **Paper set containing:**
  - Two sheets 6 in. × 6 in. French marble
  - One sheet 6 in. × 6 in. Okawara
  - One sheet 6 in. × 6 in. dark green-backed foil
  - One sheet 6 in. × 6 in. black Unryu paper
  - Two sheets 6 in. × 6 in. corrugated cardboard
  - Three sheets 12 in. × 15 in. tissue paper
  - Two sheets 12 in. × 18 in. construction paper
- **Poster of Bearden collage to be placed in room for distant view**
- **8.5 in. × 11 in. print of Bearden collage**
- **7 in. safety scissors**

Once handed out, such materials had to be collected, along with student artworks, at the conclusion of each assessment session. Lengthy scripts were prepared to make distribution and collection standardizable and feasible, and of course administrators needed to ensure that students were able to follow directions and work individually. Three-dimensional tasks that included clay tools for working with plasticine (the chosen material so schools did not have to supply water and deal with the cleanup associated with clay), various other media such as wire and cardboard, and pencil kits for sketches were that much more difficult to prepare for and clean up, after which student responses had to be recorded by photography. Further, in
some schools that had no special art room available, or even in schools that did, students did not always have the optimal space for creating.

It would have been simplest to duck the challenges posed by three-dimensional performance tasks. As Claudette Morton states in her article, *A National Arts Test*, “…[E]veryone was aware that because this assessment was so performance oriented, it was already the most expensive NAEP assessment that had ever been conducted” (Morton, 1999). Morton, a theater educator and member of the standards, framework, and assessment development committees, goes on to discuss how finally there was not enough money to assess all four arts areas at all three grades NAEP traditionally assesses. But to give up on in-depth three-dimensional creating would have meant giving up on an opportunity to explore a central component of art making, in the first large-scale, national assessment that could supply meaningful information about students’ three-dimensional abilities.

The Strategy: Separating Responding and Two-Dimensional Creating From Three-Dimensional Creating

To respond to the framework’s goals of combining assessment authenticity and depth, while attempting to mitigate administration and school burden, ETS staff and the arts development committee decided that only two-dimensional creating tasks could be combined with more than a few responding exercises. Students would complete paper-and-pencil exercises by describing, analyzing, critiquing, and interpreting works of visual arts, and then create a two-dimensional work of art linked to the responding exercises. Three-dimensional tasks that presented the additional administrative burdens discussed previously were administered separately with only a few written exercises meant to get students thinking. Even with this approach, not all students were able to take each kind of task during administration due to time constraints. Rather, some students did responding and two-dimensional creating only, and some a responding block and a three-dimensional creating task only. This arrangement had analysis implications. Still, administration took as much as 2 hours or more. It fell to Westat, the company that holds the NAEP contract for assessment administration, to work with schools to secure ample time frames, promising that the quality and engaging nature of the arts assessment would compensate for the time involved. Although this often proved to be the case, the burden of administering the assessment was considerable.

Challenges Associated With Eliciting In-Depth Responses From Students

As stated earlier, a classroom is obviously a very different setting from a drop-down assessment, for which students have had no direct preparation and during which they cannot discuss options or choices with classmates or teachers. Within the constraints of the NAEP arts blocks, the arts committee and ETS staff worked hard to prepare students for meaningful responding and creating. The arts framework states: “Because art and design activities carried out in the classroom

---

1Creating results were not summarized on a standard NAEP scale, as were responding results. To scale assessment results, there must be a sufficient number of students taking a given group of exercises, and a sufficient number of exercises to be scaled of a given type. This was not the case for the creating visual arts tasks, given administrative constraints. Because they consumed far more assessment time than did written exercises, there were fewer exercises to group together into a scale. Given the complex administrative procedures associated with these tasks, each student took only one such task. This prohibited the use of the kind of scaling methodology used to summarize responding results. Instead of a scale, creating and performing results were presented in terms of an average percentage of the maximum possible score.
27. THE NAEP ARTS ASSESSMENT

rely heavily on discussion with teachers, the assessment itself must encourage students in the absence of teachers” (NAGB, 1994, p. 27). Tasks were carefully structured to compensate—at least to a degree for the absence of classroom preparation and the fact that students’ questions about content cannot be answered during the assessment. This compensation was important for another reason. As stated previously, NAEP assessments are usually administered to random national samples of students. This means that in the case of subjects like the arts, which may not be part of a school’s regular curricula, students often have received little or no exposure to the subject. NAEP was confronted with the challenge of making assessment exercises accessible to students with little arts background yet meaningful to students with in-school exposure to the arts capable of high-level work. Short of teaching, inappropriate in an assessment, tasks were structured to create maximum context for thoughtful responding and creating.

Creating context was accomplished in several ways. All blocks, responding with two-dimensional creating or just three-dimensional creating, were organized around a single theme or problem and/or focused students on one or two stimuli. (Students were always offered high-quality reproductions as stimuli to encourage thoughtful looking. It would have been preferable to use video to display three-dimensional works, but in NAEP there were competing budgets for the theater and dance assessments, which required enormous video investments.) Building tasks around unified themes, issues, or genres gave students an opportunity to focus their attention, rather than skipping from artwork to artwork or theme to theme. This may have helped to mitigate the drop-down assessment effect and lack of time that contrasts so greatly with the kind of focused over time work that can be done in classrooms. In the two-dimensional tasks mixed with responding items, the responding items were intended to prepare students for two-dimensional creating. All tasks featured introductions to let students know what they would be doing, to supply information about the theme or artworks students would be observing, and to encourage careful work. As a rule, tasks began with simpler exercises to increase student comfort and exercises were “scaffolded” so students could address one part of a problem at a time.

Part of a responding/two-dimensional creating task for grade 8 is shown in the following section, and in its entirety at the conclusion of this chapter. In this task, or “block,” students were asked to respond to a Romare Bearden collage and then to create collages of their own. The block is designed to help students build understanding as they move through the exercises to encourage students to take the time to look carefully and repeatedly at the artwork, to encourage thoughtful responses to stimulus artworks, and to prepare them for creating work. Note the sequence of the items shown in the following. (Nancy Pistone, member of the standards, framework, and arts development committees, created the task.)

Romare Bearden Collage Task (Responding and Two-Dimensional Creating)

Take time to look at the print you have been given and the poster hung on the wall. This is a work of art by Romare Bearden called Pittsburgh Memories. Bearden is one of America’s outstanding 20th-century artists. Although he was born in North Carolina, Bearden spent time living in Harlem and in Pittsburgh. His memories of life in these places influenced his art. Pittsburgh Memories (1984) is an example of Bearden’s style of the 1970s and 1980s.

You will have 17 minutes to answer some questions about Pittsburgh Memories. You will then do a collage of your own. The questions are designed to give you an opportunity to show how well you examine and respond to this work by looking at and thinking about what you see. You should, therefore, spend time studying the work.

Consider carefully the following questions about Romare Bearden’s work, Pittsburgh Memories. Look thoughtfully at the work (your print and the poster); then write your answers as directed.
1. After you have taken some time to look at *Pittsburgh Memories*, think about your first impressions of the work. Write some words or short phrases that describe three of your first impressions. (Do NOT just state whether you think the work is good or bad, or that it is a collage.

   A

   B

   C

2. Take out a postcard, glue stick, and black felt-tip pen from your packet. Now paste your postcard in the space below. Look closely at the image. What features do you think Bearden wants you to notice in this work? Use your black felt-tip pen to draw arrows from the margin to at least three features you think Bearden wants you to notice. Label the features you have identified with brief but thoughtful descriptions, as shown in the example below.

3. Some art critics call this work a “visual narrative.” A “narrative” is an orderly account of events or a story.

   Look again at the work. What do you think is the story Bearden tells in his memory of Pittsburgh? Put into words what you think the story is about. Talk about specific things you see in the work that help you see the story.

4. Identify something in the work that is unusual or unexpected. What is it and how does it contribute to the work as a whole? Be thoughtful and specific.

Questions 5–7 are designed to help you study some of the visual characteristics of *Pittsburgh Memories*. Choose the best answer for each question.

5. Key: B

   This work by Bearden is best described as a study of
   A a still life
   B an urban landscape
   C a factory interior
   D a human figure

6. Key: B

   Which of the following is emphasized in the work?
   A The use of shading to make the subject look realistic
   B A grid-like arrangement using horizontal and vertical rectangles
   C A composition that uses traditional approaches to perspective
   D The use of a single color scheme with varying values

7. Key: D

   Which of the following most clearly identifies the style of this work?
   A Impressionism
   B Photographic realism
   C Surrealism
   D Semiabstract representation

8. Look again at the Bearden work. How does Bearden show us the contrast between the interior and the exterior areas of the building? Be specific.

Note that the task shown above begins with an introduction to situate students and encourage them to spend time looking at the work (of which two views, close and distant) are supplied. The items then move the student from an initial response, a warm-up as it were, to noting parts of the work that capture attention. The task is designed to match the way a viewer might first respond to an artwork, feeling its overall impact, while the eye takes in various parts that jump out at the viewer. The questions increase in difficulty, asking students to try and articulate the overall “story” of the work after they have noted details, and then asking them to consider the work’s technical aspects, after they have had the opportunity to explore
its affective elements. Finally, students are asked to attend to a specific device (contrast between interior and exterior space) that will have significance for the collage-making task that immediately follows. The items in the block, both the nature of the responding items and their relationship to the creating task, represent one of the finer efforts to assess what Gardner and Grunbaum (1986) call “artistic thinking.” Asking students to look carefully at artworks to explore their meanings and their forms of expression, both for preexisting works of art and those created by students in the assessment, sits at the center of the NAEP visual arts framework goals.

The importance of this goal raised concerns among members of the arts committee that students with high levels of arts knowledge and skills are not always skilled writers. Would they be able to express their responses to artworks in writing? There was discussion about other means of assessing students’ responding skills, for example, by asking students questions orally and audiotaping their answers. Apart from the fact that it is not clear that students who cannot write do better at verbalizing their thoughts, the time constraints and administrative burdens already discussed made this an unrealistic choice. Previous experience in NAEP has shown that interviewing and audiotaping students involve enormous investments of resources in equipment and staff. Further, all students would have to be assessed individually, as opposed to in classroom-sized groups. To interview individual students, and then give them suitable amounts of time to create artworks, would have demanded even more substantial administration periods than those already required by complex creating tasks. When scoring student responses, every effort was made to avoid penalizing students because they could not write well; instead, the emphasis was always on knowledge and understanding of arts content.

In spite of the challenging nature of the assessment for students and such issues as writing burden, the visual arts responding/two-dimensional creating blocks did demonstrably engage students and elicit exciting work, in the judgment of arts committee members and audiences who have attended presentations about NAEP arts. This means that NAEP did experience some success at capturing high levels of student responding and creating skills. Further, even if other students did not respond or create in skillful ways, Westat administrators repeatedly observed that students were deeply engaged in assessment tasks and teachers observing assessments taking place were quite thrilled. This in itself is important information for those interested in further arts assessment initiatives. Here are some examples of very good student work; there are two written responses and two examples of student collages, in the first case with a self-evaluation. All responses shown received the highest scores (Persky, Sandene, & Askew, 1998, pp. 102–103).

Sample Student Response Receiving a Score of Acceptable

Identify something in the work that is unusual or unexpected. What is it and how does it contribute to the work as a whole? Be thoughtful and specific.

The big train is unusual because right next to it is an apartment building and it looks like it’s going to run right into it. But the train adds more movement.
Sample Student Response Receiving a Score of Acceptable

Identify something in the work that is unusual or unexpected. What is it and how does it contribute to the work as a whole? Be thoughtful and specific.

The bright mixed colors use only one area of the picture. It shows that he is happy when he goes home. It shows his emotions change as he goes from place to place.

Look carefully at your collage.

Describe in detail the ways that you show a memory of the place you chose and explain what your collage is about. Use evidence from your work to support your answer.

In my collage, I put it as a snowstorm. There are bright green trees on the snow, and there are black and white storm clouds. I also put a cave. In the cave I have 2 people by a fire, and behind them, are 2 windows.
Blank for Self-Evaluation. Either this student ran out of time because of the effort put into the collage, or simply skipped the last question in the block. Twenty-one percent of students left this question blank, suggesting the challenge in planning assessment time for a creative, hand-on activity.

(For further samples of student works, please see the NCES/NAEP site, http://nces.ed.gov/nationsreportcard/arts/)

Some Possible Problems Associated With the NAEP Visual Arts Block Design

Although the focused blocks such as the collage task did enable relatively in-depth assessment, it limited assessment breadth in terms of both artworks and issues. It also may have inhibited performances of students not engaged by a particular problem, theme, or artwork. Some assessment designers would argue that it is preferable to mix greater numbers of stimuli and questions to enable different students more access points to assessment exercises.

Members of the development committee also worried that the organization of “respond first and then create” was too limiting. In suggesting the movement of students from creating to responding and back, the framework was seeking to allow students to bring knowledge and skills together in as flexible a way as possible. It would have been preferable to have had more creative task arrangements involving combinations of responding, two- and three-dimensional creating. Almost all creating tasks, three-dimensional included, were followed by responding items asking students to evaluate their artworks; but students often had little time to complete these questions, because test administrators were told never to interrupt students creating work to compel them to move on to the concluding questions. In any event, student responses to such questions tended to be thin, perhaps because such reflection is difficult when art making has been so rapid and constrained.

On the other hand, there was concern that the absence from three-dimensional creating of leading in-depth responding exercises may have limited students’ abilities to think through problems before undertaking such tasks. This could have been the case in spite of the context and guidance supplied by introductions and visual stimuli, and the fact that students made sketches prior to realizing three-dimensional artworks. Ultimately, any drop-down assessment in the arts (or perhaps any skills-based area) faces similar constraints as did the NAEP; creating works of art demands intellectual, emotional, and physical efforts that can be extremely difficult.
to generate on demand, and it is not possible to address every style. Given administrative
constraints, the choice seemed to be between a set of responding exercises before a two-
dimensional creating task, OR a lengthy introduction with at most two or three responding
items before a three-dimensional creating task.

Shown in the following section is an example of such an introduction and two sketching
exercises for a three-dimensional task, both intended to prepare students for creating:

In this task you will make a sculpture out of clay and wire. A sculpture is a work of art that is three-
dimensional. Three-dimensional art is not flat so you can look at it from different sides.

Sculptures can be about many different things. The sculpture you will make today will explore the
characteristics of kitchen utensils. You may have some of these utensils in your kitchen at home.

To help you think about sculptures, look at the artwork on page 3 of your booklet. [OLDENBURG]
This sculpture explores the characteristics of a pair of scissors. Even though scissors are made of hard
metal, the artist who made the sculpture used cloth and plastic to show scissors as soft and droopy.

Artists who make works like this sculpture are interested in exploring how we see and experience
objects in the world around us. Their artworks are in part about how everyday objects, like a pair of
scissors, can be seen as a kind of art.

Notice how this sculpture is surprising because it makes an everyday object into an unusual and
unexpected shape. Look at how the artist has chosen to repeat shapes and to make certain shapes large
and small. Think about how the artist may have wanted to make an artwork that would change the way
people see and think about the everyday world around them.

1. Take out the page of photographs from your packet and unfold it. Look carefully at the photographs
   of the kitchen utensils. Look at the characteristics of each utensil. Notice different shapes, details, and
   the hard materials the utensils are made of.
Your goal is to use your imagination to plan and make a sculpture out of clay and wire that will explore characteristics of the kitchen utensils, like the sculpture you just looked at explored a pair of scissors. First, you will sketch shapes and details of the kitchen utensils that interest you, then you will sketch an idea for a sculpture that combines the shapes and details you noticed, and then you will create your sculpture.

To help you begin to plan your sculpture, imagine that the kitchen utensils you see in the photograph are melting and becoming soft and bendable, so that you can make a sculpture by

- Pulling them apart
- Changing their shapes
- Combining them in different ways

A. Think about which kitchen utensils you would pull apart to use in your clay and wire sculpture. Which shapes and details of different utensils would you combine to make a sculpture?  
B. Now, right on the photographs, circle six of the shapes and details you would combine to make an imaginative sculpture.  
C. In the space on page 5, draw a sketch of each of the six shapes and details you have chosen. In your sketch, use your imagination to change the shapes, sizes, and other characteristics of the six parts by showing how they would look if they were soft and bendable.

You will have 7 minutes to draw your sketch. If you finish before time is called, you may begin work on exercise 2 on page 6. Please start.

2. Now use the sketch you just made to make a second sketch that combines the different shapes and details of the kitchen utensils you chose into an idea for a sculpture. Make your sketch of what you want your sculpture to look like in the space below.

Remember that you will make your sculpture out of clay and wire. Use your sketch to plan how you will use the clay and wire to show how shapes and details of the utensils have become soft and bendable, and to make an imaginative, standing (three-dimensional) sculpture that explores the characteristics of the utensils.

You will have 6 minutes to do your sketch.

The task shown here is meticulously designed to introduce students to a contemporary art form and approach, and to engage them in thoughtful creating. It reflects many, many iterations, and yet it is still vulnerable to criticism on at least two counts: that of creating reading burden and tasks that consist of following directions. NAEP has been accused of both, in addition to the concerns about demanding too much student writing mentioned previously. One of the great challenges in the arts assessment was finding a balance among telling students exactly what was expected of them on assessment tasks, creating context for student activities, and not requiring students to read too much. The arts committee was especially concerned about this issue given that some students who are able in the arts may not be as able at reading. Measurement experts worked closely with the arts committee to find ways of diminishing reading burden where possible. For example, every attempt was made to use simple language; complex scripts for visual arts creating tasks (as in the kitchen sculpture task) were presented to students on audiotape with which they could read along; and arts terms students might not know were defined in exercise directions, unless students’ knowledge of a particular term was being assessed (for example, sculpture and three-dimensional in the kitchen sculpture task).

Nevertheless, although further analysis would need to be done with NAEP data to prove the point, directions and introductions to arts tasks were at times lengthy and may indeed have been prejudicial for some students with strong creating abilities. Reading burden is of course not a problem in arts assessment alone. But it is especially frustrating in an assessment where other skills are so valuable.
The other question is whether student imagination and creativity are limited by tasks that include too many directions and break out too many steps. The task script shown previously tells students what to notice and takes them through a fairly analytic process of choosing shapes and forms, taking them apart, and then putting them back together. As the following example shows, some students were able to create skilled and interesting work in response to such direction (Persky et al., 1998, U.S. Department of Education, CD-ROM version). (For more examples of student kitchen sculptures, see the CD-ROM version of the 1997 NAEP Arts Report Card.)

Here is another example of task directions, this time for the collage-creating task that follows the responding items presented earlier:

Take time to look again at your print of *Pittsburgh Memories*. The collage you see visually expresses a memory of Bearden’s own past, as well as his deep appreciation for aspects of everyday life. (A collage is a work of art in which different pieces of different kinds of materials are assembled and fastened onto a flat surface.)

Study the Bearden work, and think about how the collage shows the artist’s memory of what life in Pittsburgh was like. Notice how Bearden combines and organizes objects and places in unusual and unexpected ways to express what it is like to remember. Look for the ways in which interesting contrasts between inside and outside areas and the use of details and colors communicate a memory.

Now think of a memory of a place where you once lived, where you live now, a friend’s house, or another place important to you in your community. What kinds of pictures do you see in your mind when you remember what it was like to be there?

Being as creative as you can, create a memory collage of the place you choose. In your collage, communicate what you remember about what it was like both inside of this place and outside in the neighborhood.

To make your collage:

- Take out all of the materials from your packet. You may use your scissors and/or tear materials you choose for your collage.
- Assemble on your sheet of white drawing paper pieces of any of the materials provided to show both the inside of the place you choose and what it was like outside.
- Once you have pasted down these areas, you can add details with markers and oil pastels.

In both tasks, the directions are intended to clearly present a problem to be solved, and to help students thoughtfully create. Each problem is meant to be complex enough to engage students at a high level, and their detailed presentations are intended to reach students less schooled in art making. But the very attempt to increase student comfort and create clear task parameters that would yield scorable responses may have created directions too lengthy or detailed for some students to absorb, or too inhibiting for other students very concerned with following directions. Again, further study would be needed to prove or disprove the point. At the time, arts committee members and ETS staff chose very detailed, explicit directions not only for the reasons described above but also for the purpose of successful scoring.

**BALANCING BEING OPEN AND CLOSED: ISSUES AND STRATEGIES USED IN ASSESSMENT DEVELOPMENT AND SCORING**

It seems obvious that an arts assessment ought to elicit students’ expressive abilities. The visual arts assessment tasks were all the creations of committed arts educators who greatly value...
student experimentation and expression. But assessment tasks also needed to yield scorable, comparable responses. NAEP scoring is accomplished by gathering all student responses for tasks in one central location, and assigning experts, in this case arts educators, to train small groups of qualified raters (those with suitable arts backgrounds and/or NAEP scoring experience) to assign scores to student responses. Scores are assigned to written and production work using scoring guides that feature multiple levels, or points, to accommodate a wide range of skill levels and range of responses. To train raters, experts choose sets of exemplars to demonstrate various levels of scoring criteria. As mentioned earlier, the arts committee vetted sets of exemplars to ensure that scoring matched the framework’s specifications for expected levels of skill; NAEP is not a norm-referenced assessment. Given the goal of determining whether students were demonstrating specific kinds of knowledge and skills, it was crucial that tasks express clear expectations and that scoring criteria match these expectations. There was thus a constant tension between the goals of telling students what was expected of them in creating tasks, so their responses could be meaningfully scored, and leaving enough space for students to demonstrate their expressive abilities. Scoring guides for written responses were far less complex.

Why Create Constrained, Directed Creating Tasks?

Although this book is concerned with visual arts, an example from the dance assessment is instructive. In a dance assessment task (note that dance was ultimately not assessed in 1997 due to sampling constraints), students were asked to create and perform a dance built around the theme of metamorphosis: changing from one form to another. Students were told they needed to incorporate two different shapes, levels, and movement types into their dances. The task specified the amounts of time allowed for creating and practicing the dance. Within this structure, students were free to work with their partners in any way they chose; create dances of any genre or style as long as they incorporated the criteria; and experiment with any idea of metamorphosis that engaged them.

Giving students a theme to focus on and clear specifications for their dances gave them a useful structure within which their imaginations could play, a starting point, and helped them to respond to the task. It also ensured that student dances, even if they represented different dance styles, would be clear and comparable enough to score fairly. Finally, carefully timing different stages of activities, as was also done in the kitchen sculpture task shown previously, kept students focused and engaged, and ensured that all students in the sample would have the same opportunities for creating and performing.

The priority placed on being absolutely clear and constraining in directions did generate argument among arts assessment committee members and ETS staff. Would it have been better in various cases to open up tasks to enable student demonstration of skills? For example, was it necessary that the kitchen sculpture sketching directions be so constraining as regards modes of experimentation? The conclusion at the time was that it was important to direct students and to get them working, rather than have them spend a large amount of time wrestling with what to create. And it would have been far more difficult to compare student responses for scoring purposes had students been given too many choices of theme or subject. Certainly some degree of choice was allowed. For example, students were always given two or three drawing tools and told they could use their tools in any way they chose. And, when asked to create three-dimensional works of visual art, students were not required to base their final works on their sketches. However, as a general rule, choice was limited and directions were analytic to ensure a close match between task directions and scoring criteria.
NAEP 1997 Scoring Choices for Creating Tasks

Matching task expectations and scoring criteria is a basic assessment strategy, really rule, and not news to anyone. But even with quite constrained tasks, the visual arts assessment scoring guides went through many evolutions. The rule of matching task requirements and scoring criteria is far easier to apply in an assessment where measuring more widely accepted canons of specific knowledge and agreed-upon aspects of skill are in question, as in science or math. The creation of the 1997 assessment was an occasion when all the participants were learning as they developed and scored items for a large-scale arts performance assessment.

There were two initial approaches to scoring. In one, two guides were created for a given artwork to separately address technical and expressive qualities of student artworks. The other model meticulously broke out criteria into small units based on every aspect of a problem a students was asked to solve. As an example of the latter approach, here is the three-dimensional creating task for kitchen sculpture, followed by a set of criteria for scoring. Remember that students have been asked to create two sketches prior to creating their sculptures, both of which have encouraged them to focus on showing parts of kitchen utensils as soft and bendable and combining those shapes into a plan for a sculpture.

3. Now use the clay and wire you have been given in your packet to make your sculpture. Use the sketch you have just drawn to help you, but remember that your sculpture does not have to look exactly like your sketch.
   - First work with some of your clay and/or wire to make the main shape of your sculpture.
   - Then use more of your clay and wire to add other shapes and details to your sculpture.
   - Experiment by pulling pieces of clay away from the main shape of clay and adding pieces of clay and wire to the main shape of clay.
   - Use the tools you have been given to make textures, lines, and shapes.
   - As you work, think about how your sculpture will look from different sides.

Scoring Criteria
You will have 30 minutes to create your sculpture and to answer a question about your work. The supervisor will help you keep track of the time.

Instructions: Rate each of the following items by circling the appropriate number. The numbers represent the following values:

1. Unsuccessful (did not do it)
2. Somewhat successful (partially did it)
3. Success (did it)
4. Very successful (did it very well)

How successful was the student in making a sculpture that meets each of the following criteria?

3.a 1 2 3 4 Student’s solution recognizes the properties of clay in transforming the characteristics of the objects. This is demonstrated by the student’s use of clay to show how the kitchen utensils have become soft and bendable.

3.b 1 2 3 4 Student’s solution recognizes the properties of wire in transforming the characteristics of the objects. This is demonstrated by the use of wire to show how the kitchen utensils have become soft and bendable.

NOTE: Wire can be used to show detail, and/or as an armature.

3.c 1 2 3 4 Student integrated the materials in ways appropriate to the form.
Instructions: Rate each of the following items by circling the appropriate number. The numbers represent the following values:

1. Unsuccessful (did not do it)
2. Somewhat successful (sort of did it)
3. Successful (did it)

How successful was the student in making a sculpture that meets each of the following criteria?

3.d 1 2 3 Student incorporated shapes and details of the kitchen utensils into the sculpture. (Look for more than one object, shape, and detail.)

3.e Student used the clay tools to make textures, shapes, and lines.

1 = Minimal evidence of tool use.
2 = Some evidence of tool use.
3 = Sufficient/adequate evidence of tool use.

3.f Student’s sculpture shows a variety of viewpoints or sides and is freestanding, NOT a relief sculpture.

1 = completely flat work: limited detail, one object.
2 = transitional work: some detail/elaboration, more than one object.
3 = freestanding work: fairly detailed/elaborated, several objects, includes wire.

The criteria shown here do match what students have been asked to do, but after field-testing they were felt to be too constraining. Members of the arts development committee asked one another whether we were really scoring a sculpture, a piece of artwork. Arguably, the criteria reflect the analytic directions, but even so, it was decided at the time to append an additional guide that would more holistically address the set of criteria for the sculpture:

**Holistic Guide**

The student used imagination to create an inventive, freestanding sculpture of an everyday object that is soft, bendable, or distorted. Look for evidence of ability to explore the object in the photograph to inventively combine and organize shapes, colors, details and textures into an imaginative interpretation. The materials will be used in a creative way that enables the student to effectively capture the critical relationship between the materials and the idea the student is attempting to express.

1—UNSUCCESSFUL See visual examples.
2—SOMewhat SUCCESSFUL See visual examples.
3—SUCCESSFUL See visual examples.
4—VERY SUCCESSFUL See visual examples.

Although there is some overlap between the two kinds of criteria, the goal of the second kind of guide is clearly different. It both explicitly mentions creativity and combines sculptural attributes instead of breaking them out by material used, in this case clay (in fact plasticene) and wire. Certainly the guide was easier to use than the set of more detailed criteria, and some people experienced it as a better fit with the task. But whether it was felt to be a better fit because of its references to creativity, or in fact because it attends to the fact that students were asked to explore materials to create an integrated piece, not as an end in itself, is up for discussion.
In my view, a somewhat more successful combining of criteria to evaluate solution of the problems and technical and expressive qualities is shown in the scoring guide for the collage task. This guide evolved from two guides: the one dealing with technical skill; and the other, with expressive skills. After field testing, it was felt that student responses combined technical and expressive aspects in a way that were too complex and varied to allow those qualities to be separated without overlap, that is, scoring the same thing twice. Just as the NAEP writing assessment has evolved from analytic scoring that separately evaluates grammar and development of ideas to primary trait guides that focus on a combination of prompt-related and writing skill criteria and, finally, to focused holistic guides that ask raters to evaluate the whole piece and treat the prompt purely as a springboard for writing, so many NAEP arts assessment guides ended up collapsing technical and expressive aspects into more holistic guides that score technical and expressive qualities across a number of levels, together. The arts guides did not abandon the specifics of the prompt as wholly as do the writing guides; the art “problem” students are asked to solve offers desirable parameters that enable comparisons across student responses, as demonstrated in the scoring guide for the collage task shown next.

1—UNACCEPTABLE No use of collage techniques, or very limited or ineffective use, and/or does not show inside or outside of place remembered. Objects are nonrecognizable.

2—MINIMAL The ability to use collage technique to express a memory about a place is barely evidenced in the work. Shows inside OR outside of place remembered, or shows both in a weak manner.

- Representations are of a single object or limited scheme and are simplistic and without detail.
- There is minimal exploration and variation of materials to depict objects, areas, ideas.
- A large degree of ambiguity exists because objects are incomplete; parts missing; objects float ambiguously on page.

3—UNEVEN The ability to use collage technique to express a memory about a place is evidenced in parts of the work. Shows inside OR outside of place remembered, or shows both in an uneven and unrelated manner.

- Choice and use of material show occasional awareness of pattern, texture, transparency, contrast, color, and the relationships of these qualities to depicted objects.
- Only parts of the work show exploration and variation of materials to depict objects, areas, or ideas.
- Lacks compositional unity.
- Fragments of the work evoke mood or feeling about place.

4—ADEQUATE The ability to use collage technique to express a memory about a place is evidenced throughout most of the work. Shows BOTH inside and outside of place remembered, and at least one is convincing and complex; though the two may not be well integrated, or both sides are evenly and well done (but not necessarily complex) and well integrated.

- Choice and use of material show some awareness of pattern, texture, transparency, contrast, color, and the relationships of these qualities to depicted objects.
- Forms and objects are generally shown with distinguishable features.
- Reasonably unified composition: Objects are shown in some relation to one another and to whole page.
- Work is moderately expressive.

5—EFFECTIVE The ability to use collage technique is evidenced throughout the whole work. Clearly shows BOTH inside and outside of place remembered, and the two are well integrated.
• Choice and use of material show good awareness of pattern, texture, transparency, color, and the relationships of these qualities to depicted objects.
• Material is placed in a careful, deliberate way to represent ideas.
• Forms and objects are shown with clearly distinguishable features.
• Unified composition: Objects are shown in clear relation to one another and to whole page.
• Work is very expressive.

The implied analogy made here between expressivity in art and development in writing is far from exact. Some might argue that a more apt analogy is between the creative quality in creative writing and the expressive quality in art. But it is very important to note that the language of making meaning versus that of being creative as an end in itself is closest to the arts framework. That document is most concerned with how well students can make meaning with what is available to them is a task; a concern with creativity is present in references to showing inventiveness in solutions and experimenting to seek goodness of fit with intended meaning. Creativity is seen as being embedded in experimentation and solutions that aptly express ideas, just as it is, with a very different medium; in writing. Just as the NAEP writing guides address precision in language, the ability to match word to meaning, the framework discusses using media and techniques to convey an intended meaning, and the collage guide evaluates matching collage material to idea. The distinction between successfully making meaning and being creative as an end in itself is an important distinction and fits best with a model of art education as being meaningful for many, versus a romantic model of the marginal and uniquely inspired artist.

Interestingly, it may have taken the framework writers some time to come to terms with the scoring implications of their own document, and the initial weddedness to expressivity as a separate, somehow elusive category may have been a stage in that evolution. There was much discussion about how to reward the students who seemed to demonstrate expressive skills without technical skills, but the framework does not seem to be designed to reward them. (Whether a future reworking of the framework might change this is another question.) Perhaps in this regard what is most successful about the collage guide is not its folding in of language about expressivity with technical skills, but instead its clear articulation of how well technical skills are brought to bear on making meaning. In this sense, the collage guide is very close to the NAEP holistic writing guides.

Possible Future Directions for Task Design and Scoring

A possible future approach would be to model large-scale arts assessment scoring guides even more closely on the NAEP writing assessment guides. In NAEP, students are creating what must be seen as artworks analogous to the drafty writing they create for the writing assessment. And it is precisely the fact that NAEP is an assessment meant to provide general results based on random sampling that makes it a good candidate for this holistic approach to task scoring. For example, just as writing scoring guides are organized in terms of how well students choose words, develop and organize ideas, and competently use technical skills to communicate ideas, so arts scoring criteria would reflect art educators’ agreed-upon ideas about what students at various levels of knowledge and skill ought to be able to demonstrate regardless of a specific task’s qualities. The collage guide shown previously is not too far from this, in that it takes for granted that students will apply what they may have learned about composition and use of media to create a meaningful work, just as, with only brief directions urging them to pay attention to clarity and development, students are expected to use what they know about writing in responding to NAEP writing prompts. If the collage guide were to more completely match the proposed model, it would not speak less in terms of the collage media per se and of
the goal of depicting both inside and outside of a remembered place. Instead it would focus on the use of media, in general, to create an effect, in terms of integration, clarity, and so forth.

In fact, one very good quality of the NAEP writing assessment is the room for varied student responses that may address the prompt in the most glancing fashion, or that may mix “modes,” for example, narrative and informative writing. The category of “off task” is quite rare, specifically because the writing prompts are intended to get student writing, not as ends in themselves. Instead, a specific set of writing skills is in question: word choice, development, organization, and control over conventions of grammar. In these circumstances, a student who fails to depict inside AND outside, for example, might still receive a high score, because he or she has demonstrated the ability to use arts skills to make meaning, even if that meaning is more tenuously related to the specifics of the problem students were asked to solve. (It should be noted though that students who write very “creative” narrative pieces, for example, undistinguished by technical competence, do generally receive low scores. What can be assessed and captured in a large-scale national assessment can never fulfill the promise of a small-scale, individualized, portfolio assessment, both in writing and in the arts. This is salient to the struggles mentioned previously about how to credit students who showed creative promise but seemed to lack the skill to bring that promise to fruition.)

One might even speculate as to whether arts-creating tasks could be simpler affairs, with fewer directions and steps. Such an approach would be one way to acknowledge the profoundly “draft” quality of the student artworks produced in timed assessment circumstances. Just as writing prompts, with as few words as possible, try to stimulate students to write, so arts prompts could ask students to use available tools and media to create a specific kind of two- or three-dimensional artwork, without complex directions guiding students through each step. Doing so would be less important because a wider range of student responses would be acceptable, given the more generic scoring criteria. Student responses could be standardized not in terms, for example, of whether every student attempts to depict both the inside and the outside of a place, but whether every student demonstrates suitable levels of skill in depicting anything related to the collage task. Further, as with writing planning, sketching for three-dimensional artwork would be optional, or if required, students would be told that it would not be scored, to encourage experimentation. Finally, just as the writing assessment offers students a short pamphlet on how to plan, revise, and edit their writing, so the NAEP arts assessment (or any large-scale arts assessment) could supply students with a brief pamphlet reminding them of ways to reflect on their work and its goals. This approach may enable greater student experimentation, more fluid combinations of responding and creating tasks, and simpler scoring.

Again, the analogy between arts and writing is far from perfect; it breaks down most when one considers the range of media and tools available to express art, versus writing, suggesting that art-making skills are firmly embedded in particular kinds of creating. This approach to arts task creation and scoring assumes that notions of what constitutes “good art,” or art skill, can be as agreed upon as those of “good writing.” Last, assessment developers would still be confronted with how to make creating tasks accessible to a wide range of students, and it is this need that accounted for much of the complexity and lengthy directions in the NAEP arts tasks.

But the assumption that underlies this suggestion, that there are categories of skills that can be evaluated across tasks and students, is precisely what the arts framework itself assumes in its plan for scaling student creating abilities across tasks, along with student responding abilities.

---

2Given the struggle to come to terms with appropriate skills to be assessed in writing, it seems reasonable to assume that a given generation of arts educators can reach similar consensus.
The framework assumes that the ability to bring knowledge and skills to bear upon make meaning and responding to meaning using varied media (including writing) is an ability that can be generalized from many varied tasks. Further, in the experiment carried out by Siegesmund, Diket, and McCulloch (2001), student work was scored on a range of traits regardless of whether the student precisely followed directions, suggesting that there are skills or capacities of interest to arts educators that can be scored across tasks that vary greatly in nature. (Whether the framework’s expression of such skills can be enriched with reference to the sorts of skills explored by the teachers in Siegesmund’s research is an open question. Perhaps it is these skills that NAEP committee members were trying to capture in guides about how expressive a student work is, and the next generation of the arts framework can more fully articulate this goal.)

One way to accommodate skills that run across tasks yet attend to varied kinds of arts processes might be to create two scales, one for two-dimensional creating and one three-dimensional creating, and retain the 1997 NAEP model in which responding was treated separately. This might enable a more sensitive portrait of arts knowledge and skills; the peculiarity of tools and media in the arts is such that it is challenging to assess student skill without profoundly relating that skill to its expression with particular media and tools. Or, perhaps the scales are more properly divided by categories that capture the sensory, unique quality of art, as described by Siegesmund et al. (2001).

CONCLUSION

The previous discussion seeks to articulate several tensions in the creation and scoring of the NAEP arts assessment: making tasks feasible for administration yet authentic in the terms presented in the NAEP arts framework; encouraging thoughtful student responding and creating work without burdening students with too many directions and constraints; enabling students from a wide range of arts backgrounds to perform on the assessment, again without undue reading burden or constraints; and enabling student responses to be scorable without making tasks too limiting.

Clearly, certain compromises had to be made to successfully administer an arts performance assessment of a random national sample of students. Responding items to assess students skills at analyzing and articulating how artworks convey meaning required students to respond in writing, because interviewing individual students was not a possibility. Such items, designed also to move students thoughtfully toward creating activities, could not be used to precede three-dimensional creating tasks, because of the complexities inherent in the administration of such tasks. To compensate for lack of context created by the absence of responding items and to engage students of varying levels of exposure to arts education, directions and introductions to three-dimensional tasks were rather lengthy and complex. And, both two- and three-dimensional creating tasks limited, to some degree, student choice of problem, approach to solution, and media to ensure that students could complete assessment work in a limited time frame and to make tasks standardizable so student responses could be scored.

Concerns about the compromises made to create the NAEP arts assessment can be responded to in several ways. One way is to discuss the reasons behind the choices made, as this chapter does. The dilemmas of administrative and school burdens, limited budget, and reaching students of widely varying skill levels are not limited to NAEP, and are likely to be confronted by anyone seeking to assess large numbers of students in a drop-down assessment, even at the state or local level. Another response is to note that the assessment did yield results that many many
people in the field of arts education have learned from. Although researchers and educators such as Siegesmund et al. (2001) are critical of some of the approaches NAEP took, they are nevertheless using the NAEP model as a starting point for future assessment design in the arts. Also, students enthusiastically wrote about and created artworks for the NAEP, and assessment results showed a range of skills from the most elementary grasp to the most sophisticated abilities. And, positive correlations were noted in the 1997 arts report card among various in (and out of school) arts activities and performances on both responding and creating tasks. (Persky et al., 1998, pp. 104–115). These correlations were sustained and further explored by secondary research, for example, that performed by Read M. Diket (2000). Finally, one can make suggestions for the future that may mitigate some of the problems of the assessment, such as more holistic scoring guides and associated simpler tasks.

Still, it is clear from the remarks above that success at assessing arts at a national level is intrinsically limited. The challenges posed by a large-scale administration are real. Further, it is inevitable, even appropriate, that an assessment of such large proportions, aimed at a random national sample of students, will offer a quite general measure of students’ arts knowledge and skills. Again, in this respect, the NAEP writing assessment is similar; who would argue that the assessment offers the same picture of student skill as a portfolio assessment of students’ works over several months, or the observations of a classroom teacher? But that general measure may be a fine snapshot of the state of arts education at a given time. State and local assessments may lose in budget what they gain in time and flexibility of administration. By demonstrating what can be achieved on a national level, the NAEP plays an extraordinarily important role in defining the outer limits of what is possible in large-scale arts assessment. The suggestions about holistic approaches to tasks and scoring offered earlier may make this role even more possible in the future.

**TASK APPENDIX**

Romare Bearden Collage Task (Responding and Two-Dimensional Creating)

Take time to look at the print you have been given and at the poster hung on the wall. This is a work of art by Romare Bearden called *Pittsburgh Memories*. Bearden is one of America’s outstanding 20th-century artists. Although he was born in North Carolina, Bearden spent time living in Harlem and in Pittsburgh. His memories of life in these places influenced his art. *Pittsburgh Memories* (1984) is an example of Bearden’s style of the 1970s and 1980s.

You will have 17 minutes to answer some questions about *Pittsburgh Memories*. You will then do a collage of your own. The questions are designed to give you an opportunity to show how well you examine and respond to this work by looking at and thinking about what you see. You should, therefore, spend time studying the work.

Consider carefully the following questions about Romare Bearden’s work, *Pittsburgh Memories*. Look thoughtfully at the work (your print and the poster), then write your answers as directed.

1. After you have taken some time to look at *Pittsburgh Memories*, think about your first impressions of the work. Write some words or short phrases that describe three of your first impressions. (Do NOT just state whether you think the work is good or bad, or that it is a collage.)

   A 
   B 
   C
2. Take out a postcard, glue stick, and black felt-tip pen from your packet. Now paste your postcard in the space below. Look closely at the image. What features do you think Bearden wants you to notice in this work? Use your black felt-tip pen to draw arrows from the margin to at least three features you think Bearden wants you to notice. Label the features you have identified with brief but thoughtful descriptions, as shown in the example below.

3. Some art critics call this work a “visual narrative.” A “narrative” is an orderly account of events or a story.

   Look again at the work. What do you think is the story Bearden tells in his memory of Pittsburgh? Put into words what you think the story is about. Talk about specific things you see in the work that help you see the story.

4. Identify something in the work that is unusual or unexpected. What is it and how does it contribute to the work as a whole? Be thoughtful and specific.

Questions 5–7 are designed to help you study some of the visual characteristics of *Pittsburgh Memories*. Choose the best answer for each question.

5. Key: B

   This work by Bearden is best described as a study of
   
   A a still life
   B an urban landscape
   C a factory interior
   D a human figure

6. Key: B

   Which of the following is emphasized in the work?
   
   A The use of shading to make the subject look realistic
   B A grid-like arrangement using horizontal and vertical rectangles
   C A composition that uses traditional approaches to perspective
   D The use of a single color scheme with varying values

7. Key: D

   Which of the following most clearly identifies the style of this work?
   
   A Impressionism
   B Photographic realism
   C Surrealism
   D Semiabstract representation

8. Look again at the Bearden work. How does Bearden show us the contrast between the interior and the exterior areas of the building? Be specific.

9. Take time to look again at your print of *Pittsburgh Memories*. The collage you see visually expresses a memory of Bearden’s own past, as well as his deep appreciation for aspects of everyday life. (A collage is a work of art in which different pieces of different kinds of materials are assembled and fastened onto a flat surface.)

   Study the Bearden work, and think about how the collage shows the artist’s memory of what life in Pittsburgh was like. Notice how Bearden combines and organizes objects and places in unusual and unexpected ways to express what it is like to remember. Look for the ways in which interesting contrasts between inside and outside areas and the use of details and colors communicate a memory.

   Now think of a memory of a place where you once lived, where you live now, a friend’s house, or another place important to you in your community. What kinds of pictures do you see in your mind when you remember what it was like to be there?

   Being as creative as you can, create a memory collage of the place you choose. In your collage, communicate what you remember about what it was like both inside of this place and outside in the neighborhood.
To make your collage:

- Take out all of the materials from your packet. You may use your scissors and/or tear materials you choose for your collage.
- Assemble on your sheet of white drawing paper pieces of any of the materials provided to show both the inside of the place you choose and what it was like outside.
- Once you have pasted down these areas, you can add details with markers and oil pastels.

After you have completed your memory collage, you will answer a question about your collage. You will have 43 minutes to do your collage and to answer the question.

10. Look carefully at your collage.

Describe in detail the ways that you show a memory of the place you chose, and explain what your collage is about. Use evidence from your work to support your answer.

Kitchen Sculpture Task (Three-Dimensional)

In this task you will make a sculpture out of clay and wire. A sculpture is a work of art that is three-dimensional. Three-dimensional art is not flat, so you can look at it from different sides.

Sculptures can be about many different things. The sculpture you will make today will explore the characteristics of kitchen utensils. You may have some of these utensils in your kitchen at home.

To help you think about sculptures, look at the artwork on page 3 of your booklet. This sculpture explores the characteristics of a pair of scissors. Even though scissors are made of hard metal, the artist who made the sculpture used cloth and plastic to show scissors as soft and droopy.

Artists who make works like this sculpture are interested in exploring how we see and experience objects in the world around us. Their artworks are in part about how everyday objects, like a pair of scissors, can be seen as a kind of art.

Notice how this sculpture is surprising because it makes an everyday object into an unusual and unexpected shape. Look at how the artist has chosen to repeat shapes and to make certain shapes large and small. Think about how the artist may have wanted to make an artwork that would change the way people see and think about the everyday world around them.

1. Take out the page of photographs from your packet and unfold it. Look carefully at the photographs of the kitchen utensils. Look at the characteristics of each utensil. Notice different shapes, details, and the hard materials the utensils are made of.

Your goal is to use your imagination to plan and make a sculpture out of clay and wire that will explore characteristics of the kitchen utensils, like the sculpture you just looked at explored a pair of scissors. First, you will sketch shapes and details of the kitchen utensils that interest you, then you will sketch an idea for a sculpture that combines the shapes and details you noticed, and then you will create your sculpture.

To help you begin to plan your sculpture, imagine that the kitchen utensils you see in the photograph are melting and becoming soft and bendable, so that you can make a sculpture by

- Pulling them apart
- Changing their shapes
- Combining them in different ways

A Think about which kitchen utensils you would pull apart to use in your clay and wire sculpture. Which shapes and details of different utensils would you combine to make a sculpture?

B Now, right on the photographs, circle six of the shapes and details you would combine to make an imaginative sculpture.
C In the space on page 5, draw a sketch of each of the six shapes and details you have chosen. In your sketch, use your imagination to change the shapes, sizes, and other characteristics of the six parts by showing how they would look if they were soft and bendable.

You will have 7 minutes to draw your sketch. If you finish before time is called, you may begin work on exercise 2 on page 6. Please start.

**Scoring Criteria**

1—UNACCEPTABLE Characteristics in the sketch are not made to appear bendable, soft, or distorted. Or, student only shows one changed characteristic.

2—PARTIAL Characteristics in the sketch are made to appear bendable, soft, or distorted in minimal ways; little experimentation is in evidence. Or, student only shows two changed characteristics.

3—ESSENTIAL Characteristics in the sketch are made to appear bendable, soft, or distorted; some elaboration of details and experimentation is in evidence. Or, student only shows three or four changed characteristics.

4—EXTENDED Five or six characteristics in the sketch are made to appear bendable, soft, or distorted. There is strong evidence of elaboration and experimentation with shapes and details.

**NOTE:** Whereas simply eliminating a portion of a tool is change, showing distortion of that tool is experimentation. Experimentation of size, shape, and line can be understood as distortion of the following kinds:

- Elongation
- Integration
- Simplification
- Elaboration
- Abbreviation
- Stylization
- Juxtaposition
- Transformation/metamorphosis

2. Now use the sketch you just made to make a second sketch that combines the different shapes and details of the kitchen utensils you chose into an idea for a sculpture. Make your sketch of what you want your sculpture to look like in the space below.

Remember that you will make your sculpture out of clay and wire. Use your sketch to plan how you will use the clay and wire to show how shapes and details of the utensils have become soft and bendable, and to make an imaginative, standing (three-dimensional) sculpture that explores the characteristics of the utensils.

You will have 6 minutes to do your sketch.

**Scoring Criteria**

1—UNACCEPTABLE Sketch shows objects as they are. There is no evidence of variation and experimentation. No elaboration of details.

2—INSUFFICIENT Objects show limited evidence of variation and experimentation. Characteristics are made to appear bendable, soft, or distorted in minimal ways. Little or no elaboration of details. Images are ambiguous.

3—UNEVEN Objects show uneven or inconsistent evidence of variation and experimentation and are combined in predictable ways. Characteristics show evidence of appearing bendable, soft, or distorted. Minimal elaboration of details.

4—ADEQUATE Objects show sufficient evidence of variation and experimentation and are combined
in fairly inventive ways. Characteristics show clear evidence of appearing bendable, soft, or distorted. There is clear evidence of elaboration of details.

5—ELABORATED Objects show considerable evidence of variation and experimentation and are combined inventively. Characteristics show strong evidence of appearing bendable, soft, or distorted. Details are well elaborated.

NOTE: Variation and experimentation of size, shape, and line can be understood as distortion of the following kinds:

- Elongation
- Integration
- Simplification
- Elaboration
- Abbreviation
- Stylization
- Juxtaposition
- Attenuation
- Transformation/metamorphosis
- QUANTITY of variation is less important than interesting contrasts and juxtapositions. Avoid a prejudice toward complexity.
- Whisks, unless the drawing is exceptional, are scored 1.

3. Now use the clay and wire you have been given in your packet to make your sculpture. Use the sketch you have just drawn to help you, but remember that your sculpture does not have to look exactly like your sketch.

First work with some of your clay and/or wire to make the main shape of your sculpture.

Then use more of your clay and wire to add other shapes and details to your sculpture.

Experiment by pulling pieces of clay away from the main shape of clay and adding pieces of clay and wire to the main shape of clay.

Use the tools you have been given to make textures, lines, and shapes.

As you work, think about how your sculpture will look from different sides.

You will have 30 minutes to create your sculpture and to answer a question about your work. The supervisor will help you keep track of the time.

**Scoring Criteria**

**Instructions:** Rate each of the following items by circling the appropriate number. The numbers represent the following values:

1. Unsuccessful (did not do it)
2. Somewhat successful (partially did it)
3. Success (did it)
4. Very successful (did it very well)

How successful was the student in making a sculpture that meets each of the following criteria?

3.a 1 2 3 4  Student’s solution recognizes the properties of clay in transforming the characteristics of the objects. This is demonstrated by the student’s use of clay to show how the kitchen utensils have become soft and bendable.

3.b 1 2 3 4  Student’s solution recognizes the properties of wire in transforming the characteristics of the objects. This is demonstrated by the use of wire to show how the kitchen utensils have become soft and bendable.
NOTE: Wire can be used to show detail and/or as an armature.

3.c 1 2 3 4 Student integrated the materials in ways appropriate to the form.

**Instructions:** Rate each of the following items by circling the appropriate number. The numbers represent the following values:

1. Unsuccessful (did not do it)
2. Somewhat successful (sort of did it)
3. Successful (did it)

How successful was the student in making a sculpture that meets each of the following criteria?

3.f 1 2 3 Student incorporated shapes and details of the kitchen utensils into the sculpture. (Look for more than one object, shape, and detail.)

3.g Student used the clay tools to make textures, shapes, and lines.

1 = Minimal evidence of tool use.
2 = Some evidence of tool use.
3 = Sufficient/adequate evidence of tool use.

3.h Student’s sculpture shows a variety of viewpoints or sides and is free-standing, NOT a relief sculpture.

1 = Completely flat work: limited detail, one object.
2 = Transitional work: some detail/elaboration, more than one object.
3 = Free-standing work: fairly detailed/elaborated, several objects, includes wire.

**HOLISTIC GUIDE:** The student used imagination to create an inventive, free-standing sculpture of an everyday object that is soft, bendable, or distorted. Look for evidence of ability to explore the object in the photograph to inventively combine and organize shapes, colors, details, and textures into an imaginative interpretation. The materials will be used in a creative way that enables the student to effectively capture the critical relationship between the materials and the idea the student is attempting to express.

1—UNSUCCESSFUL See visual examples.
2—SOMEWHAT SUCCESSFUL See visual examples.
3—SUCCESSFUL See visual examples.
4—VERY SUCCESSFUL See visual examples.

Look carefully at the sculpture you have made, and imagine that your sculpture is going to be part of an art exhibit at your school.

What would you want people looking at your sculpture to notice and think about? Be specific.

Explain why.

**REFERENCES**


27. THE NAEP ARTS ASSESSMENT


