

Drawing on Education: Using Drawings to Document Schooling and Support Change

WALT HANEY
MICHAEL RUSSELL
DAMIAN BEBELL
Boston College

In this article, Walt Haney, Michael Russell, and Damian Bebell summarize a decade of work using student drawings as a way to both document and change education and schooling. After a brief summary of more than one hundred years of literature on children's drawings, the authors point out that drawings have been little recognized as a medium of educational research in recent decades. Next they explain how the work reported here has evolved, recounting how they have used student drawings as a way to document educational phenomena. They then present reliability and validity evidence to support such use on a macro level. The authors go on to relate examples at the micro level of how drawings have been used to inform and change education and learning. Finally, they argue that student drawings, though only one form of inquiry, help illustrate the fundamental point that, if educational reforms are to succeed, we must treat teachers and students not just as the objects, but also as the agents, of reform and improvement.

Drawings have been used for decades as markers and mirrors of personal identity. . . . Drawings offer a different glimpse into human sense making than written or spoken texts do because they can express that which is not easily put into words: the ineffable, the elusive, the not-yet-thought-through, the subconscious. . . . Children have control over drawing, which for them is a natural form of symbolic expression. The notion of child-in-control is a challenge to the power that adults hold over children (Grugeon, 1993). Adults are usually considered to have an advantage over children in areas of oral and written expression, but few adults in Western societies develop a visual vocabulary and drawing skills much beyond an elementary school level, probably because drawing is not as highly

valued as reading or writing (Wilson & Wilson, 1977). Thus, in print-oriented cultures, drawing in a sense puts children on a more equal footing with adults in terms of adequacy of expression. If we wish to know more intimately what children think and feel, we might begin by taking their drawings more seriously. Although art educators have been aware of this for decades, most educational research has not. (Weber & Mitchell, 1995, pp. 34–35)

Introduction

In times of war and crisis, children's drawings have frequently been recognized as offering a unique window on events and their meaning — not just for children, but also for adults. In the wake of the 9/11 attack on the United States, for example, a number of observers sought to use drawings to help document the horrific events of that day and people's reactions to them. In one effort, the New York University Child Study Center and the Museum of the City of New York collected drawings and paintings of the events surrounding 9/11 by children between the ages of five and eighteen from New York, New Jersey, and Connecticut. A selection of these drawings has recently been published in *The Day Our World Changed: Children's Art of 9/11* (Goodwin & Fahstock, 2002). As a *New York Times* book reviewer wrote of this volume, "Sometimes it takes a child's-eye view to transcend artifice" (Heller, 2002, p. 11).

In this article, we summarize a decade of research, by ourselves and our colleagues, on drawings about education. We argue that drawings by children, adolescents, and adults of both epic events like those of 9/11 and everyday situations involving classrooms, schools, and learning have unusual power to document and change the educational ecology of classrooms and schools. We argue, as do Weber and Mitchell (1995), that while many have recognized the power of children's drawings for more than a century, educational research generally has not paid serious attention to this topic.

Perhaps the most obvious technique employed to document what goes on in schools and classrooms is observation. When performed carefully, observation requires that one or more people regularly visit the school or classroom. To gain insight that is more than superficial and impressionistic, observers must be trained. To obtain results that are generalizable, observation must occur over an extended period of time. For these reasons, systematic observation of schools or classrooms is often quite expensive. Moreover, observation, unless part of intensive ethnographic inquiry, may fail to illuminate the realities of life in classrooms from the perspective of students. A considerable body of research, ranging from *Pygmalion in the Classroom* (Rosenthal & Jacobson, 1968) to Claude Steele's (1997) work on how stereotype vulnerability can affect the achievement of females and ethnic minorities, shows that it is not just how students are educated, but also *the expectations and stereotypes surrounding their education* that can affect their learning. This means it is vital to inquire into students' own conceptions of their educational experiences and learning.

FIGURE 1 *Drawing by a Middle School Student of a Teacher at Work in a Restructured Classroom*



The basic idea behind this research was to develop an alternative method to document the lives of schools, classrooms, and students by drawing, quite literally, on the insights and perspectives of those who are perhaps the most assiduous observers of school and classroom life, namely, students. Specifically, we sought to develop methods for documenting what goes on in schools, classrooms, and in the student learning process by asking students to draw pictures. As we will explain, in response to carefully constructed prompts, surprising aspects of school and classroom activity and student learning can be documented via student drawings. Furthermore, drawings can provide a valuable catalyst to document, change, and improve what goes on in schools. In short, although the research reported here will by no means settle many of the controversies and problems currently swirling around educational assessment and school reform, the methods of using student drawings that we describe hold considerable promise for making assessment more useful to schools, teachers, and students, and for documenting and broadening educational reform efforts.

Our initial interest in student drawings grew out of work begun under our Urban District Assessment Consortium project in 1992 and was extended over three years in connection with the Co-NECT school design. The Co-NECT model of school reform emphasized, among other things, intensive use of technology and project-based learning.¹ Boston College's Center for the Study of Testing, Evaluation, and Educational Policy (CSTEAP) helped Co-NECT schools implement a model of school accountability that employs matrix sampling of multiple forms of assessment (including multiple-choice, writing, and performance assessment) from different representative samples

of students. In the spring of 1994, as part of a broader set of assessments aimed at measuring students' attitudes toward school, we began asking students in one Co-NECT school to draw pictures of one of their teachers at work in the classroom. In 1995, we then used this drawing prompt at five additional Co-NECT schools. We found that these drawings were powerful in two respects. First, they seemed particularly useful in promoting reflection by teachers regarding their methods. Second, we found that analysis of these drawings was a simple but powerful way to document changes in the educational ecology of schools.

During the mid-1990s, we also provided technical assistance to urban school districts funded under the Edna McConnell Clark Foundation's Program for Student Achievement. As a result of our introduction of the idea of student surveys, two school systems (Corpus Christi, Texas, and San Diego) undertook student reflection surveys in 1996 and 1997 in all their middle schools (12 schools in Corpus Christi and 22 in San Diego). These surveys included a student drawing exercise. We found that we were able to analyze reliably the features appearing in student drawings and to use these analyses to document changes in schools over time. In the first-year surveys, in response to the prompt asking them to draw a picture of one of their teachers at work in the classroom, most students in most schools drew a picture of a teacher-directed classroom, often with no student even appearing in the drawing (see Figure 2).

However, in analyzing results from spring 1997, we found that in some schools a substantially larger proportion of drawings depicted students working together (see Figure 3).

Equally important, the student drawings proved to be a powerful vehicle for teachers to learn from students' perspectives. For example, we received the

FIGURE 2 *Middle School Student's Drawing of a Teacher-Directed Classroom*

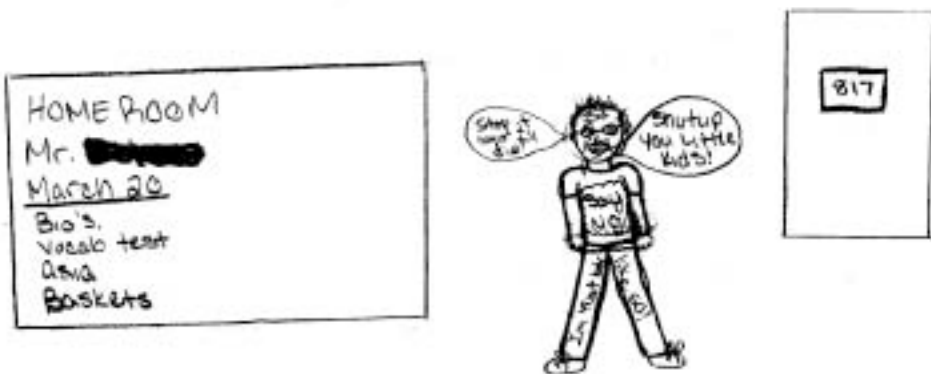
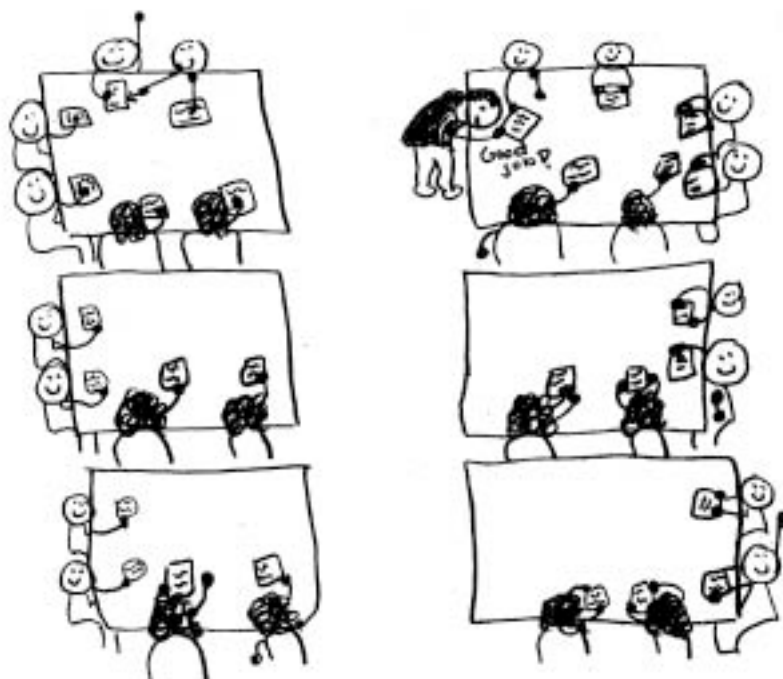


FIGURE 3 *Middle School Student's Drawing of a Cooperative Classroom*



following report on the reflection surveys from a Corpus Christi Co-NECT project coordinator:

Several district principals and teachers stated that the survey had provided the most insightful information they had received regarding student perceptions of the academic standards and the strategies utilized by teachers, or the lack of such, in teaching the standards.

Middle school principals shared individual school reports with staff in whole school faculty groups or by team/house in an informal study group format, and teachers used the rubric developed by CSTEPP to analyze student drawings depicting a teacher at work in the classroom. Analysis of the drawings and use of the rubric helped teachers to see the classroom through the eyes of students who “tell it like it is.” (P. Lyons, personal communication, July 23, 1997)

Other testimony to the value of drawings comes from a middle school in which we helped introduce drawings as part of the Co-NECT project some years ago. At the Scott Middle School in Hammond, Indiana, which had been working toward full inclusion of special needs students, principal Frank Lentvorsky reported that using drawings was “especially helpful in allowing us to learn the feelings of our learning disabled students.” He added that “middle school kids in general have a difficult time communicating with adults. . . .

The drawing lets us get some of the nonverbal communicators to give us their impressions” (Tovey, 1996, p. 6).

Finally, thanks to a major research grant from the Spencer Foundation, we and our colleagues have been able to gather a wide range of evidence on the reliability and validity of using drawings to make inferences about classrooms and schools, and even state policies concerning testing. It is worth noting that our research with drawings has been both fundamental basic research and collaborative action research. In terms of the former, our inquiries have sought to document the validity of student drawings as indicators of the educational ecology of classrooms and schools. As anyone familiar with the evolution of educational evaluation over the last forty years knows, documenting the implementation of educational programs or reforms has been a major difficulty (Haney, 1977; Pressman & Wildavsky, 1973). The research reported here aimed, in part, to study the validity with which student drawings reflect what occurs in classrooms and schools. If such drawings do provide valid indicators of classroom processes, student drawings of their teachers, classrooms, and learning experiences may offer a way to document educational processes that is far easier to employ widely than are systematic observations.

But more important than traditional validity concerns may be the value of such drawings in terms of what Samuel Messick (1989) calls consequential validity, that is, the power of assessment to effect change and improvement. Our work in collaboration with a number of teachers and schools has led us to conclude that student drawings are not only a simple and powerful way of documenting the educational ecology of classrooms and schools, but also a potentially powerful tool for reflection and change on the part of teachers and students. Our use of drawings in both basic and applied research is, we admit, somewhat unusual. But on this point we agree with Ellen Lagemann (1996) who, citing John Dewey, observes that “education requires scholarship that erodes common boundaries between research, development and practice” (p. 20). Our work concerning drawings grew out of collaborative efforts with schools to evaluate their own reform efforts, evolved into basic research to document the reliability and validity of drawings as a form of assessment, and, as we explain later in this article, is rebounding back to practice, concerning not just assessment and evaluation, but also teaching and learning.

Children’s Drawings

As Claire Golomb notes, “For over a hundred years, the drawings of children have enchanted a rather diverse audience of psychologists, educators, art historians and artists” (1992, p. 1). As early as 1885, Ebenezer Cooke published an article on children’s drawings in which he described the successive stages of development as he had observed in them. He urged that “art instruction in the schools be made to conform more nearly to the mentality and interests of the child” (cited in Goodenough, 1926, p. 1). Since then, the study of chil-

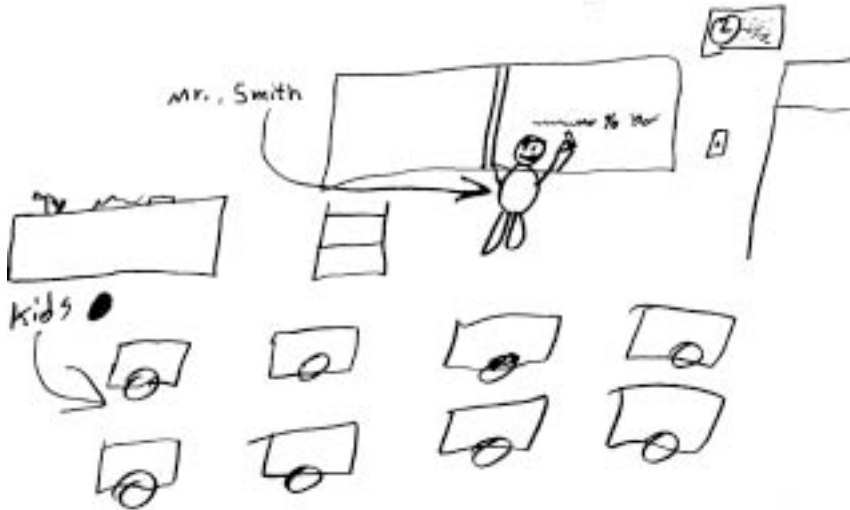
dren's drawings has attracted the attention of scholars ranging from Sigmund Freud and Jean Piaget to Rudolph Arnheim and Howard Gardner. In many of his inquiries into children's social and spiritual lives, Robert Coles has drawn on children's drawings. Interest in children and adults' drawings of children has continued in fields as diverse as art education, clinical psychology, and even intelligence testing. Florence Goodenough's classic book, *Measurement of Intelligence by Drawings* (1926), gave rise to the Draw-A-Man test, which in the 1980s continued to be one of the ten most widely used psychological tests in clinical practice.²

There is not space here to summarize much of the scholarly work on children's drawings over the last century, but we note four broad patterns in this vast literature. Psychological perspectives that view drawings in light of children's cognitive development or emotional concerns have dominated most of this literature. With only rare exceptions have drawings been used to illuminate the educational contexts in which students find themselves (Weber & Mitchell, 1995). Yet as Golomb (1992) notes, "Any notion that children create [drawings] in a social vacuum is quite untenable" (p. 5). Studies of drawings by victims of the Holocaust, by children in cities beset by violence, and by children of divorce have made this fact abundantly clear. Yet we have been able to locate only a handful of studies that use drawings as a vehicle for examining students' understanding of classrooms or schooling. These include work in Toronto in the 1960s on children's drawings of their classrooms by Rogers (1969) and others; Black's 1991 examination of college students' drawing of themselves writing a paper for school; Gamradt and Staples' 1994 study of children's drawings of their schools; and Weber and Mitchell's 1995 study of images of teachers.

Another pattern reveals that most of the literature on drawing deals with young children, rather than with older children, adolescents, or adults. Virtually all of the drawings discussed in Golomb's (1992) review, for example, were made by children under age twelve. Similarly, Goodenough's (1926) development of the Draw-A-Man test was based on samples of children aged four to ten, and this test continues to be more widely used with children than adults (except for mentally retarded or verbally incapacitated adults). There are a few instances in which drawings by older subjects have been studied. These include Siegel's (1986) use of drawings by adults who were terminally ill with cancer; Golomb's (1992) review of drawings by adolescent victims of the Holocaust; Branfman's (1972) compilation of drawings by adults who were victims of the bombing of villages in Laos during the war in Indochina. But as far as we have been able to learn, there have been few studies of drawings by students across the full range of schooling, and even fewer studies have focused on students' depictions of schooling and learning.³

A third pattern shows that, as part of increased interest and use of performance assessment (Haney & Madaus, 1989; Rudner & Boston, 1994), drawings have only recently been used in large-scale research. For example, the

FIGURE 4 Drawing of a Traditional Classroom with Students at Desks



Third International Math and Science Study (TIMSS) made far more use of performance assessments than any previous international comparative educational study. One item used to assess students' science achievement in the middle school years asked, "Draw a diagram to show how the water that falls as rain in one place may come from another place that is far away" (Beaton et al., 1996, p. 62). Though only this one drawing item was included in the TIMSS science assessments, its inclusion demonstrates that drawing exercises can be used in large-scale, cross-cultural research (over 40 countries participated in TIMSS, with data collected in more than 30 languages) and be reliably scored.

Our fourth pattern is that drawings are rarely used in educational research. Despite the century-old tradition of using children's drawings in psychological research, very little educational research, other than that focused on art education, has employed drawings. One way to illustrate this general pattern is by searching databases such as the Educational Resources Information Clearinghouse (ERIC) using search terms such as *children's art* and *standardized tests*. (Houston, 1986/1990). In the vast literature covered in the ERIC system, far more has been written about standardized tests than about children's art. Another example of the neglect of drawings in educational research is found in the second edition of *Complementary Methods for Research in Education* (Jaeger, 1997), which provides a wide-ranging review of educational research methods, both qualitative and quantitative. Although this volume includes a section on arts-based educational research, it deals only with textual narrative forms of art-based inquiry and does not mention using drawings as a method of educational research.⁴

FIGURE 5 *Elementary School Student's Drawing of a Teacher Using a Computer*



Yet, as Weber and Mitchell (1995) suggest, if educational researchers want to know more about what children think and feel about their school experiences and to give them more active control over their learning, a good place to begin is to take children's drawings more seriously. This is exactly what we have done over the last decade — use student drawings as a tool of educational research in both large-scale and classroom-based action research.⁵

Over the last decade, we found that asking students to draw a picture of one of their teachers at work in the classroom is a useful way to document changes in classrooms undergoing restructuring. It has also proved a powerful means of helping teachers to reflect on and think about changing their classroom practice. We initially started using student drawings in spring 1994, as part of school accountability assessments in the Co-NECT project.⁶ Our approach involved having random samples of students complete multiple-choice, open-ended, and performance assessments in reading, writing, mathematics, and science. We also asked them to draw a picture in response to the following prompt: "Think about the teachers and the kinds of things you do in your classrooms. Draw a picture of one of your teachers working in his or her classroom."⁷ The prompt was chosen after several were pilot tested, and was designed to elicit youngsters' images of their classroom environments and learning activities, not to focus on particular teachers.

After the first set of drawings was collected, we developed a check-list to indicate whether they showed particular features. For instance, were the teachers depicted alone or with students? Were they addressing the class or assigning homework? What subject were they teaching, and were computers shown in the drawings? After checking the reliability with which two different raters

FIGURE 6 *Elementary School Student's Drawing of a Teacher with a Circle of Students*



coded features of the drawings, we proceeded to code drawings from five different schools in 1994 and 1995.⁸

At schools that had just begun to implement the Co-NECT model, students overwhelmingly depicted their teachers at work in front of the entire class, often at the blackboard or seated behind a desk. Most showed the teacher working alone. If students were present at all, they were most often seated in rows, as in Figure 4, which shows a drawing in which faceless and even bodiless students are represented as circles at desks. We found the presence of the clock — which shows up frequently in students' drawings — to be an indication that the clock remains a defining technology of education (Mumford, 1934).

After a year or more, significantly larger numbers of children drew pictures showing students using computers and working in groups (both features of the Co-NECT design). Figure 5 indicates the motivational power of introducing elementary children to computer use. In this instance, the teacher was in control of the technology with students merely observing. In several other drawings, however, after just one school year of ready access to computer technology, students depicted themselves at the keyboard with the teacher nearby as cheerleader and coach.

Figure 6 shows another kind of drawing that was increasingly common after one year of restructuring via the Co-NECT project. Here we see not only the teacher working with students, but also students with faces and even smiles.

In the next two parts of this article, we describe how we have used drawings to document educational phenomena, to establish the reliability and validity of these methods, and to help *inform and change* education and learning. In the conclusion we discuss practical aspects of working with drawings, as well as some limitations.

Using Drawings to Document Educational Phenomena

Our work with student drawings began as a way to document the educational environments of classrooms and schools from students' perspectives. Over the years we experimented with several prompts and various approaches to summarize the patterns apparent in sets of drawings. Our prompts have included:

1. Think about the teachers and the kinds of things you do in your classrooms. Draw a picture of one of your teachers working in his or her classroom.
2. Think about all of the different things your teachers do with you in the classroom. Draw a picture of what a camera would see when one of your teachers is working in the classroom.
3. Think about the kinds of work and activities you do in your classes. Draw a picture of what a camera would see when you are learning in the classroom.
4. Think about all of the different things you do when you read. Draw a picture of what a camera would see when you are reading.
5. Think about the steps you take when writing a paper for school. In the space below, draw a picture or series of pictures that reflect your writing process.
6. Think about the teachers and the kinds of things you have done in your class today. Draw a picture of your teacher teaching and yourself learning.
7. Draw a picture of yourself taking the MCAS (Massachusetts Comprehensive Assessment System, the high-stakes test introduced in Massachusetts in 1997).
8. Think about the math work and activities you do outside of school. Draw a picture of yourself learning math outside of school.
9. Think about the work and activities you do in math class. Draw a picture of yourself learning math in school.

With most of these prompts, we collected sets of drawings and then sought to identify patterns in them. We focused on identifying patterns in sets of drawings rather than on the meaning behind individual drawings, because absent the opportunity to talk with the artist, the meaning of any drawing may remain unclear. Golomb (1992) provides several examples in which doubtful, if

not downright incorrect, psychological interpretations have been made of individual children's drawings.

The following experience illustrates the danger of reading too much into any one drawing. In the winter of 1996, during a visit to the Barton Open School in Minneapolis, Minnesota, one of the authors was asked to model how to elicit student drawings and discussions of them. In one middle school social studies classroom he asked students to think about what their teachers had been doing in their classes and to draw a picture of one teacher at work in the classroom. As often is the case, students became deeply engaged in creating their drawings and had to be coaxed into completing them. Upon review, we found that two students had depicted a most unusual scene: an image of their science teacher standing in front of the blackboard with flames coming out of his pocket. This unusual scene might readily provoke a variety of interpretations, but what became apparent through discussion was that both students were depicting an incident that had actually happened some months before. The teacher in question had a nervous habit of fiddling with change in his pocket as he taught. Once, while fiddling with the contents of his pants pocket, two books of matches had rubbed together and caught fire. We return to this example later when we discuss the extent to which student drawings represent typical circumstances they have actually experienced, their stereotypes of teachers and classrooms, or unusual or funny experiences that stand out in their memories.

We have used four different approaches to identifying patterns in sets of drawings: emergent analytic coding, trait coding, holistic coding, and holistic review.⁹ For the first three approaches we have documented the reliability of coding and, as we explain, in several instances we investigated the validity of inferences drawn from these methods.

Emergent Analytic Coding

For sets of drawings elicited in response to most of the prompts listed above, we use emergent analytic coding to develop a checklist of features the drawings may contain. To develop this checklist, two investigators independently review a sample of about fifty drawings and record various features. In drawings of teachers or classrooms, for example, features that may be coded include whether a blackboard or clock is depicted and whether "teacher talk" is represented. The two checklists are then compared and condensed into a list of features to be used as a draft-coding sheet. This draft-coding sheet is used by two raters to code a second sample of about fifty drawings. The two raters work independently and code features as either present or absent. In addition, the raters make note of features that are present in the drawings but absent from the coding sheet. The coding results for the two raters are then compared. Formal descriptions are developed of features that have high levels of agreement. Investigators work together to examine drawings for which there are discrepancies, to identify reasons for those discrepancies, and to finally de-

velop a definition of the feature. If a common definition of the feature cannot be developed, the feature is removed from the list. Similarly, if the coders note additional features not on the original list, these features are defined and added. Once formal descriptions for all features are developed, the list is used by a new pair of raters to independently code a third sample of drawings. Their ratings are then compared. Features that show low levels of reliability (less than 80% agreement) are removed from the list. The remaining features and their accompanying descriptions are then used to score the full set of drawings.

Trait Coding

In the analytic or checklist approach just described, we seek to document whether specific features are present in drawings. In trait coding, we analyze drawings at a higher level of abstraction and rate them with regard to the extent to which a particular trait is represented. The following examples illustrate two applications of trait coding.

To help estimate if and when high school students write papers for school using computers, Russell (1999), one of the authors of this article, asked students to draw a picture of their writing process. To help interpret the drawings, students were also asked to describe them. Then the drawing and accompanying descriptions were reviewed to develop the following coding scheme, the extent to which computers were used in students' writing:

- 0 — Blank
- 1 — Computer not visible
- 2 — Computer visible during final draft
- 3 — Computer visible during editing
- 4 — Computer prominent throughout the writing process

When this coding scheme was applied, it became apparent that the two middle categories — 2, computer visible during the final draft and 3, computer visible during editing — could not be reliably distinguished. In addition, all of the students had participated in the drawing activity, so the "Blank" rating category could be dropped. Hence the coding scheme was condensed to:

- 0 — Computer not visible
- 1 — Student writes on paper and then works on computer near the end of writing process
- 2 — Computer prominent throughout the writing process

The condensed coding scheme was then used to code the full set of student drawings, with results that will be described in more detail below.

Another example of trait coding comes from Gulek's (1999) multi-trait, multi-method study of modes of classroom instruction. In this study, over 350 students in grades 3–6 in twenty classrooms were asked the following: "Think

about the teachers and the kinds of things you have done in your class today. Draw a picture of your teacher teaching and yourself learning.” To code the drawings, Gulek developed a four-point scale to rate the extent to which drawings showed:

- 1 — Highly teacher-directed mode of instruction
- 2 — Moderately teacher-directed mode of instruction
- 3 — Moderately student-centered mode of instruction
- 4 — Highly student-centered mode of instruction

For each point on the scale, a list of features associated with the level of student-centered or teacher-directed instruction and sample drawings was prepared to create a guide for rating the mode of instruction represented in students’ drawings. For example, the features and sample drawings characterizing “Highly teacher-directed mode of instruction” were:

Score of 1 (Highly teacher-directed mode of instruction):

- Only the teacher depicted, students are not present in the picture.
- If depicted, student desks are in rows.
- The teacher is depicted at the blackboard, or at teacher’s desk.
- Teacher talk, if any, is lecturing or disciplining.

After a check on the reliability with which the guide could be applied, it was used to code Gulek’s full set of drawings. We discuss the results of our reliability analysis below, after we recount two other approaches to the interpretation of patterns in drawings.

Holistic Coding

In this approach, drawings are coded using a four- or five-point scale, based on an overall judgment about how some overall or holistic aspect of a situation is depicted. For example, using this approach, drawings of classrooms have been rated in terms of the extent to which a learning environment is depicted positively or negatively. Another example is that students’ drawings of their processes of writing a paper might be rated in terms of the complexity of their understanding — that is, their metacognition — of their own writing processes.

Holistic Review

The three approaches to identifying patterns in sets of drawings described above are fairly time consuming and labor intensive to develop and apply. For this reason and, more importantly, as a way to engage teachers and others in using the results of drawing exercises, we have also developed what we call a process of holistic review of drawings. In this approach we ask groups of three teachers (or sometimes students) to review a set of thirty to fifty drawings and to answer three questions:

FIGURE 7 Score 1 Sample Drawings



1. What patterns do you see in the drawings?
2. Why do you think these patterns occur?
3. What do you think might be done differently in your school as a result of what you see in the drawings?

The groups typically spend 20–25 minutes on this assignment and often report their findings back to the larger group. The patterns identified frequently correspond with the features documented by the analytic or trait coding. In addition, when reviewing drawings by their own students, teachers almost always can identify aspects of their classrooms that were missed by outside raters. In the next section we recount more of our experience in using drawings with both teachers and students.

Evidence of Reliability and Validity

We have examined several lines of evidence on the reliability and validity of using drawings to make inferences about aspects of education. This sort of inquiry is important because scholars have recently pointed out that, despite the popularity of the clinical use of drawings (e.g., to assess personality traits, emotional states, and even intelligence or developmental levels), there is little empirical evidence of the reliability of such interpretations (Thomas & Jolley, 1998).

We have pursued two kinds of reliability studies related to research using students' drawings. The first focuses on the reliability with which drawings can be coded. In essence, coding reliability is akin to inter-rater reliability. The second addresses the reliability or stability of the information gleaned from drawings over time. In this case, reliability is akin to test-retest reliability.

Coding Reliability

Depending on the type of evidence sought from student drawings, one of four methods of coding or review described above has been employed. For all three coding approaches, we examined inter-rater reliability using one of three methods: correlation between ratings, percent agreement, and Cohen's kappa.

For particular sets of drawings, two raters were employed to code a subset of fifty drawings. Table 1 summarizes the correlations between the ratings by two independent raters for each of four prompts. For drawings that were coded using an analytic approach, the median correlation is presented. On average, we found a high level of inter-rater reliability for the analytic coding. Holistic and trait coding also yielded relatively high levels of inter-rater reliability.

We also examine the percent agreement between raters; that is, the percentage of cases in which two independent raters agreed in their ratings of a set of drawings. Table 2 summarizes the percent agreement for drawings elic-

TABLE 1 *Summary of Inter-Rater Correlations of Ratings of Drawings Elicited with Different Prompts*

| <i>Prompt</i> | <i>Coding Method</i> | <i>Inter-Rater Correlation</i> |
|--|----------------------|--------------------------------|
| Think about all of the different things your teachers do with you in the classroom. Draw a picture of what a camera would see when one of your teachers is working in the classroom. | Analytic | .86 |
| Think about all of the different things you do when you read. Draw a picture of what a camera would see when you are reading. | Analytic | .88 |
| Think about all of the different things you do when you read. Draw a picture of what a camera would see when you are reading. | Holistic | .68 |
| Think about the kinds of work and activities you do in your classes. Draw a picture of what a camera would see when you are learning in the classroom. | Holistic | .82 |
| Think about the steps you take when writing a paper for school. In the space below, draw a picture or series of pictures that reflect your writing process. | Trait | 1.00 |
| Think about the teachers and the kinds of things you have done in your class today. Draw a picture of your teacher teaching and yourself learning. | Trait | .84 |

Note: All correlations statistically significant at the 0.95 level of confidence.

ited with several different prompts. For analytic coding results, the percent agreement between two raters across all features is presented. For holistic and trait ratings, the percent agreement is in ratings between two raters across ratings. Across these prompts and the other two coding methods, there are reasonably high levels of agreement between rater codes.

Note that the correlation and percent agreement results are generally lower for the holistic codes than for the analytic codes. This may occur in part because agreement for the analytic codes can be exaggerated due to the infrequency with which some features appear in the drawings. In other words, agreement may occur simply because a feature is not present in many drawings. A simple example may help illustrate why simple “percent agreement” can be a misleading measure of reliability of classifications when one is dealing with skewed distributions. Suppose two raters, A and B, independently rate a set of ten student drawings of teachers at work in their classrooms with

TABLE 2 Summary of Percent Agreement across Ratings of Drawings Elicited with Different Prompts

| Prompt | Coding Method | Percent Agreement |
|--|---------------|-------------------|
| Think about all of the different things your teachers do with you in the classroom. Draw a picture of what a camera would see when one of your teachers is working in the classroom. | Analytic | 94 |
| Think about all of the different things you do when you read. Draw a picture of what a camera would see when you are reading. | Analytic | 89 |
| Draw a picture of yourself taking the MCAS. | Analytic | 90 |
| Think about all of the different things you do when you read. Draw a picture of what a camera would see when you are reading. | Holistic | 76 |
| Think about the kinds of work and activities you do in your classes. Draw a picture of what a camera would see when you are learning in the classroom. | Holistic | 79 |
| Think about the steps you take when writing a paper for school. In the space below, draw a picture or series of pictures that reflect your writing process. | Trait | 100 |

regard to whether or not the teacher is shown with a smile on his or her face, coded with a 0 if no smile is apparent, and a 1 if a smile is depicted. Let us suppose further that in eight of ten drawings no smile is apparent. However, on the last two drawings it is ambiguous whether a smile is depicted. Suppose that rater A codes drawing nine with a 1 and ten with a 0, but that rater B does the reverse. If we simply calculated percent agreement across the two sets of ten ratings, we would find 80 percent agreement (8 of 10 drawings were coded 0 by both raters), even though the two raters completely disagreed as to which drawings did show a smile. The need to take into account the likelihood of chance agreement when dealing with skewed distributions has long been recognized. The best-known statistic for taking these factors into account is Cohen's kappa (Cohen, 1960).

Without trying here to explain technical details of how this statistic is calculated, we note simply that Cohen's kappa coefficient may be interpreted as the proportion of consistent classifications *beyond* those expected by chance. Table 3 summarizes Cohen's kappa results across independent ratings of drawings elicited with three different prompts. As can be seen, the frequency with which raters agree remains much larger than by chance alone. In his method-

TABLE 3 *Summary of Cohen's Kappa Results for Independent Ratings of Three Sets of Drawings*

| <i>Prompt</i> | <i>Coding Method</i> | <i>Cohen's Kappa</i> |
|--|----------------------|----------------------|
| Think about all of the different things your teachers do with you in the classroom. Draw a picture of what a camera would see when one of your teachers is working in the classroom. | Analytic | 0.78 |
| Think about all of the different things you do when you read. Draw a picture of what a camera would see when you are reading. | Analytic | 0.68 |
| Draw a picture of yourself taking the MCAS. | Analytic | 0.67 |

ological note on Cohen's kappa in *Psychological Reports*, Kvalseth (1989) suggests that a kappa coefficient of 0.61 represents "reasonably good" overall agreement. By this standard, our methods for rating of drawings have been shown to be more than "reasonably good."

In sum, our inquiries via correlational analyses, calculation of percent agreement, and Cohen's kappa provide evidence that drawings produced in response to a variety of prompts can be coded reliably.

Stability of Drawing Results

Another aspect of reliability often distinguished from inter-rater reliability is stability. Stability refers to the extent to which assessment results are stable over time. Members of our group have investigated this issue in two quite different ways. In his research, Gulek (1999) used drawings from students to provide information about the mode of instruction used by their teachers. Here the unit of analysis was the classroom. To examine the reliability of information provided by drawings over time, Gulek administered the drawing prompt mentioned previously twice in each of three classes, with about four weeks between administrations. All of the drawings were then rated using the same trait-coding guide pertaining to class mode of instruction. Averaging across all drawings for each classroom, a mean trait score was calculated for each classroom, for each occasion. The mean trait scores for classrooms were then compared to examine their stability over time. Finding that the trait scores on drawings collected four weeks apart were similar, Gulek concluded that for this trait, mode of instruction, evidence supported the stability of information provided by drawings over time.

Another issue relating to the stability of drawing results over time is the extent to which students' drawings of their classrooms or teachers represent re-

cent or typical classroom experience, on the one hand, or some funny or memorable event. Recall, for example, the incident from the Barton school in Minneapolis, when two students drew pictures of their science teacher standing in front of the blackboard with flames coming out of his pocket — an incident that had actually happened several months prior. In another incident, during a holistic review of drawings with teachers, one teacher was surprised to see that several students had depicted an incident in which she had spilled her coffee cup, with liquid dripping down the front of her desk, even though the incident had happened many weeks before.

Such experiences have led us to seek additional evidence on the degree to which drawings represent typical classroom experience. In a study with middle school students, Bebell (2001) gathered direct evidence on this issue. Nearly five hundred drawings were collected from sixth-, seventh-, and eighth-grade students in the Watertown, Massachusetts, public schools in response to the prompt, “Think about the teachers and the kinds of things you do in your classrooms. Draw a picture of one of your teachers working in his or her classroom.” After students completed their drawings, they were asked to fill out a short questionnaire, which asked, among other things, whether the students’ drawings showed “A teacher you like,” “A teacher you dislike,” “Your favorite teacher,” “Your favorite class,” “What your classes are generally like,” and/or “A funny or memorable experience.” A summary of results is shown in Table 4.

A preponderance of students (75%) reported drawing a teacher they like, while a sizable minority (19%) reported drawing a teacher they disliked. Nearly two-thirds (65%) said they drew what their classes are generally like, but nearly a third indicated they had depicted a funny or memorable incident. Bebell (2001) reports that results regarding what was drawn did not vary much by gender, but his findings do suggest that about one-third of students drew some sort of critical incident from the past. This clearly suggests that unusual or funny events in the lives of classrooms or schools may well affect the stability of results of drawing exercises over time.

Validity Studies

Through our work with different schools and research initiatives, we have collected four types of validity evidence. As part of his work on computers and writing, Russell (1999) provided evidence of the construct validity of information collected via drawings. Gulek’s (1999) multi-method, multi-trait study also provided construct evidence. Finally, our work with teachers and that of Sack (1997), as well as cases in which people have adapted drawing exercises for their own use, provide evidence that information provided from drawings can have positive consequences for teachers and their classroom practices. In other words, this latter kind of evidence could be construed as a kind of consequential validity evidence for drawings.

TABLE 4 *Results of Survey of Watertown, Massachusetts, Middle School Students Regarding What They Drew (N= 491)*

| <i>Did you draw . . .</i> | <i>Percent Yes</i> |
|---------------------------------------|------------------------|
| a teacher you like? | 74.6 |
| a teacher you dislike? | 18.8 |
| your favorite teacher? | 41.9 |
| your favorite class? | 41.2 |
| what your classes are generally like? | 65.2 |
| a funny or memorable experience? | 31.8 |

— Construct Validity/Factor Analysis

As part of a larger study on the effects of administering tests via computer, a survey was used to collect information about students' prior computer use, particularly for writing (Russell, 1999). In addition to several closed-ended questions about students' computer use, the survey contained a drawing prompt asking students to draw their writing process, as mentioned previously. In addition to the survey, students also completed a keyboarding test to document their typing skill, or, more specifically, their keyboarding speed. Together, the data from these instruments provide two types of evidence that drawings can be used to provide valid information regarding students' computer use.

First, to examine the dimensions measured by the student survey, a principal components analysis was performed. Russell (1999) identified three components through this analysis. The first component dealt with students' reported use of computers in their writing process and evidence of computer use in their drawings of the writing process. The second component was defined by use of a computer at home and the length of time students have used a computer. The third component was less well defined, but was related to students' preference for using computers and use of computer in school. Together these three accounted for 54.8 percent of the variance in the responses to the student questionnaire. Particularly revealing was that the coding of the drawing prompt (developed to collect information about when computers enter students' writing process) fell into the first component. This provides empirical evidence that the students' drawings were consistent with their responses to the survey.

In addition, the relationship between the information provided by the drawing and students' keyboarding scores provides additional evidence that the drawings elicited information about similar characteristics. The key-

boarding test provided a measure of how quickly a student could use a keyboard to enter text into a computer. The drawing prompt was designed to provide a measure of how much experience a student had working with a computer when writing papers. One would expect that the higher a student's level of computer use when writing a paper, the more skilled she or he would be at using a keyboard. To examine the relationship between these two measures, we examined bivariate correlations between students' keyboarding scores measured in words per minute (WPM) and the trait ratings of their drawings. Although the correlation between these two measures was relatively low (.20), it was positive and significant at the .99 level of confidence.

— Construct Validity/Convergent Evidence

Gulek (1999) used three methods of inquiry (a student survey, student drawings, and videotapes of classrooms) to study two aspects or traits of classroom practice, namely, the mode of instruction and the variety of learning materials used. The surveys contained closed-ended items and two scales were produced, one for mode of instruction and one for variety of materials. As described above, the drawings were also coded using trait scoring, with respect to depiction of both aspects or traits of classroom practice. To code the videos, Gulek employed time sampling and used the same trait-coding guide he used with the drawings. The data were then summarized for each of the twenty classrooms included in his study. For his study, the classroom was the unit of analysis.

To examine the construct validity of drawings using a multi-trait, multi-method inquiry (Campbell & Fiske, 1959), Gulek (1999) calculated the correlations between the scores yielded by each method of assessment for each of the two traits. This was done for all twenty classrooms studied. The fundamental logic of multi-trait, multi-method inquiry is that, if particular methods of assessment yield valid indicators of the traits of interest, then two different methods of assessing the same trait ought to correlate more highly than measures of two (presumably uncorrelated) traits using the same method. These two aspects of multi-trait, multi-method inquiry are referred to as the demonstration of both convergent validity (two different methods of measuring the same trait will yield results that converge) and discriminant validity (one method of measuring two different traits will show results that do not correlate) (Campbell & Fiske, 1959). With regard to discriminant validity, Gulek found that for each of the three methods of inquiry, results showed essentially zero correlation between the two traits measured (the survey results showed a correlation between the mode of instruction and variety of learning materials traits of 0.13; the drawing results showed a correlation of 0.16 between the two traits; and the video results showed a correlation of -0.21). Results across methods for the two traits are shown in Table 5. For mode of instruction, results from drawings correlated significantly with both survey results ($r = 0.52$) and video results ($r = 0.58$). Results were slightly less clear for the variety of

TABLE 5 *Inter-Correlations for Mode of Instruction and Variety of Materials Traits (N = 20 classrooms)*

| | <i>Survey</i> | <i>Drawing</i> | <i>Video</i> |
|-----------------------------|---------------|----------------|--------------|
| <i>Mode of Instruction</i> | | | |
| Survey | 1.00 | | |
| Drawing | .52* | 1.00 | |
| Video | .27 | .58** | 1.00 |
| <i>Variety of Materials</i> | | | |
| Survey | 1.00 | | |
| Drawing | .25 | 1.00 | |
| Video | -.29 | .40 | 1.00 |

**p<.01 * p<.05

learning materials trait. In both cases, results from the drawings correlated more highly with the video results than with the survey results. Since the videos provide a more accurate picture of what actually occurs in the classroom, the stronger relationship between the drawing and the video results provides evidence that drawings provide valid information on what actually occurs in the classroom.

In sum, we have collected a range of reliability and validity evidence to show that drawings can be used in a variety of ways to document real and important aspects of classroom environments and student practice. In terms of reliability, we have shown that using analytic, trait, or holistic coding approaches, independent raters can reliably code features of traits depicted in drawings. We have also shown that drawings done four weeks apart can be used to document stable aspects of classroom environments — though we also have evidence to indicate that funny, unusual, or memorable events in the lives of classrooms may also affect what shows up in student drawings. We have also summarized three lines of validity evidence indicating that drawings can be used to document real aspects of classroom life and students' practice. Russell's (1999) study regarding computers and writing showed that features appearing in students' drawings of their own writing processes correlated with direct student reports about their use of computers in writing. The study also showed that trait ratings of how prominently computer use appeared in drawings correlated with performance tests of students' keyboarding skills. Gulek's (1999) multi-trait, multi-method inquiry into the ecology of classrooms provides both convergent validity evidence and discriminant validity evidence supporting use of drawings as a means of documenting mode of

classroom instruction. The fourth and final form of validity evidence — regarding what might be termed consequential validity — is described in the next section of this article.

Using Drawings to Inform and Change Education and Learning

Our research indicates that the systematic coding of drawings can validly and reliably document change in the organization of classrooms. Yet, from the first time we tried this method (using the holistic review approach) with teachers, we were most struck by the power of drawings to engage teachers and provoke reflection and change. We began this use of drawings in our work with the Co-NECT project. After conducting multiple forms of assessment (e.g., multiple-choice tests, written tests and performance assessments, surveys of student attitudes and drawings) in Co-NECT schools, we met with school staff to present and discuss the results. From the first time we used student drawings at the Accelerated Learning Laboratory School in Worcester, Massachusetts (1994), it was apparent how engaging the drawings were for the teachers. We presented random samples of fifty drawings by students in the school to groups of three or four teachers. The teachers were asked to flip through the drawings and look for patterns, speculate about their causes, and think about what they might do differently, based on the drawings. The first staff development session lasted well beyond its scheduled end because teachers were so interested in reviewing and talking about the drawings.

When teachers examined the results of the multiple-choice, open-ended, performance assessment and survey items, discussion typically focused on whether or not different subject-matter areas needed more emphasis in their teaching. Teachers often raised the issue of whether or not the content of the assessment matched the school's curriculum. When results were disappointing, it was common for someone to suggest that the cause might be a mismatch between what was tested and what was taught. Although teachers had been engaged in their school's reform efforts, discussion of student assessment results tended to focus on how well different subjects were being taught — discussions that generally proceeded with little if any reference to the school's reform efforts.

However, after reviewing the drawings (often after some joking and laughter about how teachers' physical features had been drawn), discussion turned to the teachers themselves, and not just to what was being taught, but how. The drawings drew teachers into exploring questions such as how they could spend less time at the blackboard and more time with the students; how they might structure their classrooms differently so as to encourage students to focus more on each other and less on the teacher; and how the teachers could integrate more cooperative activities into their classroom. In short, the drawings proved an effective way to focus teachers' attention on how they were

teaching, how students were engaged in the classroom, and how the reform efforts were affecting their classroom teaching.

How Drawings Have Supported Change in Other Contexts

In addition to our own experiences described above, we have seen — in work by ourselves, our students, and our colleagues — that drawings can support action research and change in a variety of contexts.

— Use of Drawings with Student Teachers

Sack's (1997) research showed that drawings could be used effectively with a range of beginning teachers. In this work, Sack asked five full-time elementary student teachers from five different teacher-education institutions to draw pictures of themselves working in the classroom. Simultaneously, their cooperating teachers and their students were asked to draw pictures of their "student-teacher" teachers at work in the classroom. Sack found that having student teachers review these drawings provided a powerful means for student teachers to learn about their own practice. She concluded that "the use of drawings coupled with indirect interviews [represents] an attractive alternative . . . method in fostering reflective and self-aware teachers" (p. 208).

— Use of Drawings in Promoting Reflective Reading

In another study, Lifford et al. (2000) found that drawings could be used effectively not just with teachers, but also with students. Lifford and her English teacher colleagues at Dedham (MA) High School asked their students to think about all the steps they went through from the time they began to read to the time they felt they had made sense of what they read, and to "try to capture all that is involved in a series of drawings or pictograms" (p. 51). These educators found that "not only do these drawings provide us with a great deal of insight into how our students read, but they have also helped the students themselves to understand in a most concrete way some of the thinking that goes on in their brains when they read" (p. 51).

— Other Examples

There are a number of other examples of cases in which educators have used drawings to provide insight into educational practices and issues, not as formal research endeavors but as *in situ* efforts at educational improvement. We know of two college professors, for example, who regularly use drawing exercises as part of their normal course evaluations. Similarly, we have worked with an elementary school principal in New Jersey who asked students in her school to "draw a picture of your principal working" as a form of professional self-evaluation. Additionally, since publishing our study of students' drawings of themselves taking the Massachusetts MCAS test (Wheelock, Bebell, & Haney, 2000), we have been contacted by people in two other states who want

to use the drawing approach to examine how students are dealing with their states' tests. Even more recently, groups of teachers in the Needham and the Wellesley, Massachusetts, public schools have undertaken studies of students' math learning by having students draw pictures of themselves learning math (Arora, Berger, & Young, 2002).

These examples are hardly definitive proof of the power of drawings to help promote reflection and change in teachers, schools, and students. At the same time, in our extensive experience with many forms of educational research, it is unprecedented for teachers and others to seek out spontaneously a new way of inquiring into educational practice and learning on the basis of simply reading a short article. Thus, the examples presented here provide clear signs of the potential value of using student drawings as a means of documenting the educational ecology of schools and classrooms. These examples also show how student drawings can promote action research and teacher reflection. If nothing else, we have repeatedly seen the power of such drawings, if not to transcend artifice altogether, at least to delve beneath easy assumptions about the lives of schools, classrooms, and students.

Conclusion

Classroom research examining teaching and learning, present and past, still remains sparse in this country. Few researchers are willing to invest the large blocks of time necessary to make sense of the complexity of classrooms. Moreover, the sources of data are heavily oriented to adults — teacher interviews, direct observation of teaching, documents written about and by teachers, etc. Securing data from students is most difficult. Student perspectives on what is happening in a classroom are seldom explored by researchers. Frederick Erickson and Jeffrey Schultz wrote a piece in the *Handbook of Research on Curriculum* a few years ago, pointing out how this has been a seriously underdeveloped line of inquiry about both teaching and learning. That was in 1992. Few researchers have taken up the challenge. (L. Cuban, personal communication, September 12, 1997)

The work reported in this article might be viewed as an effort to address the problems described by Larry Cuban in 1997. For nearly a decade we have explored the use of student drawings of their teachers, classrooms, and learning experiences as a way to examine educational life in schools from the perspective of students — a perspective too widely ignored. Today, the need for classroom research of this sort seems even greater than it was a decade ago. The successive waves of education “reform” since the *Nation at Risk* report in 1983 have only accelerated, culminating in the No Child Left Behind Act of 2001. This act and many similar reform proposals have emphasized external mandates for schools, stressing policies such as high-stakes testing, and rewards and sanctions for so-called high- and low-performing schools and teachers. Such external mandates for improving schools are, we think, of limited value

FIGURE 8 *Drawing by First-Grade Student Epitomizing the Teacher's Status with Students*



unless greater attention is paid to what goes on behind classroom doors and in the minds of students. Yet remarkably, even though most recent education reform proposals have been predicated on promoting the future well-being of our nation's young people, as Cuban (1997) pointed out, "student perspectives on what is happening in classroom" are seldom explored by researchers, much less policymakers.

Student drawings provide a rich opportunity to document students' perspectives and to transcend assumptions and artifice regarding what is going on in classrooms. Yet though literature on children's drawings has accumulated in the fields of clinical psychology, art education, and child development, drawings have been largely neglected as a tool of educational research.

In this article, we have built the case that our approaches yield reliable and valid evidence from drawings. With regard to validity, results of students' drawings of their writing process corresponded with survey question results about their use of computers at school and at home. Furthermore, results of multi-trait, multi-method analysis yielded evidence of both convergent and discriminant validity. We have also shown that student drawings have consequential validity, in that they seem to promote reflection and change, not via formal research but via action research, or even teaching. In contrast, the stability of results for describing classrooms over time is less clear. Specifically, we have

documented the fact that results of parallel drawing activities carried out in the same classrooms a month apart are similar, but that funny, memorable, or traumatic events can affect how students depict their classrooms, making student drawings less stable over time.

In closing, we offer one general disclaimer, some practical advice, and a final thought about using drawings as a method of educational research and reflective educational practice. First, while we have focused in this article largely on drawings, it is by no means because we do not value other methods of research. Indeed, in general we are fans of mixed or complementary methods of educational research and inquiry (Jaeger, 1997). If one relies on just one lens of inquiry to view any educational (or other) phenomena, one can never know the extent to which the lens itself distorts what is seen. Thus, though we focused on research on drawings here, we advocate using drawings in concert with other methods of research and inquiry.

Second, as the “science teacher with flames coming out of his pocket” example nicely illustrates, it will inevitably be impossible to interpret the meaning of some drawings without some access to the artists’ thoughts as to what was drawn. When working with relatively small numbers of students or teachers, it is probably best to interview the artists about what was drawn, as Sack (1997) and Lifford et al. (2000) were able to do in their studies. However, when working with much larger numbers of drawings, this may be a practical impossibility. Nonetheless, we have learned that, whenever possible, it can be a great aid to interpretation and coding to gather from the artist a short description of the drawing, or at least a checklist explicating what was drawn.

Third, with regard to practical advice, our thinking regarding the seemingly simple choice of soliciting drawings done in black and white or in color has evolved. Early in our work, because of the prohibitive costs of reproducing color drawings for review by teachers with whom we were working, we decided to stick with black-and-white drawings.¹⁰ More recently, however, as the price of color reproduction has fallen and the possibility of including color illustrations in scholarly journals on the Internet have increased, we have done more with color drawings (e.g., Wheelock, Bebell, & Haney, 2000).

Another practical matter deals with people’s attitudes toward drawing. As Weber and Mitchell (1995) observed, drawing is for children “a natural form of symbolic expression” (p. 35). However, we have found that, at least in the United States, by the middle teenage years or so, some students will resist engaging in drawing exercises, sometimes commenting, “I can’t draw.” This is, we think, a remarkable example of a “learned disability,” for clearly all people can and have drawn as children. But by the middle teenage years or later some people have been “taught” that they cannot or are not good at drawing. Interestingly, this may be a somewhat culturally specific phenomenon. There is evidence that in some countries such as China and Japan, drawing seems to be a more universally maintained and valued skill than it is in the United States

(Deguchi, 1998; Winner, 1989). It is important to assure people that the exercise is not a “test” of their drawing skill, but rather an alternative way of documenting and making visible their thinking and feelings about X (where X is the focus of the drawing exercise).

Our work with drawings has surely only begun to scratch the surface of how they can be used as tools for educational research, inquiry, and teaching. We readily acknowledge that, in an article-length publication, we cannot do justice to the previous literature on children’s drawings. In closing, let us simply offer a slight explanation of the title of this article and of the drawing reproduced in Figure 8. The phrase “drawing on education” is of course a minor play on words. What we have tried to recount here is not just the value of using drawings about education and learning, but also the value of “drawing out” education. It should be recalled that, as the *Oxford English Dictionary* tells us, the modern word “educate” derives from the Latin words *educare*, to rear, bring up (children, young animals), and *educere*, to lead forth. The drawing reproduced in Figure 8, rendered by Seth Lavenski of his first-grade teacher in 1998, symbolizes what we have seen again and again in our work with student drawings of their teachers, classrooms, and learning. Such student drawings serve to make visible and remind us of the incredible standing and influence teachers and education have in the lives of young people — usually, but not always, for good.

Notes

1. The Co-NECT model was developed by Bolt Beranek and Newman (BBN), with funding from the New American Schools Development Corporation (NASDC). More information on the Co-NECT model is available at <http://www.co-nect.com/>.
2. Lubin, Larsen, and Matarazzo (1984) also report that, in their 1982 survey of clinical psychologists, the House-Tree-Person drawing test was also among the top ten tests used.
3. Work in Toronto in the 1960s on children’s drawings of their classrooms is a notable exception.
4. The authors of this section, Tom Barone and Elliot Eisner, point out that “most arts-based educational inquirers have, at least up to this time, employed words as their medium of expression” (Jaeger, 1997, p. 73).
5. We have found the power of drawings so valuable that Walt Haney has begun using drawings as a way to gain insight into his own teaching. Specifically, as part of an effort to evaluate his own graduate teaching, he asks students to draw a picture of him teaching. See Haney’s website at www2.bc.edu/~haney.
6. Haney used drawings in research with refugees in Laos in the early 1970s and with graduate students studying alternative forms of inquiry about technology in the 1980s, but it was not until 1994 that drawings were used in one of CSTEOP’s research projects.
7. This prompt was developed after experimentation with a number of variants. We found, for example, that if we asked simply, “Draw a picture of one of your teachers,” some students would draw just a profile.
8. It is worth noting that, because of some controversy in the past about psychological interpretations of children’s drawings, in our analysis we sought to focus our initial coding scheme on surface features like those mentioned.

9. As far as we know, none of these approaches to coding drawings has previously been used to make sense of patterns in sets of drawings. Nonetheless, we acknowledge that these approaches are directly analogous to approaches widely used to analyze text. The analytic or checklist coding is analogous to word frequency analyses of text, which date back to the early 1900; see, for example, the work of Edward Lee Thorndike (Clifford, 1968). The trait and holistic coding approaches are analogous to trait and holistic ratings of essays (Applebee, 1984, 1994). Similarly, the holistic review approach is analogous to focus group reviews (Krueger, 1994).
10. Black pen is better than pencil in our experience. The latter can smudge when drawings are handled a lot. Also be sure to supply good plain white paper because relying on paper that people happen to have with them will result in a hodgepodge of styles and sizes of paper.

References

- Applebee, A. (1984). Writing and reasoning. *Review of Educational Research*, 54, 577–596.
- Applebee, A. (1994). English language arts assessment: Lessons from the past. *English Journal*, 83(4), 40–46.
- Arora, A., Berger, D., & Young, J. (2002). *Student drawings: A study of the Black and White performance gap in a suburban elementary school system*. Unpublished manuscript.
- Beaton, A., Martin, M., Mullis, I., Gonzalez, E., Smith T., & Kelly, D. (1996). *Science achievement in the middle school years: IEA's Third International Mathematics and Science Study (TIMSS)*. Chestnut Hill, MA: Boston College TIMSS, International Study Center.
- Bebell, D. (2001, April). *1,000 words: What can we learn from middle school students' classroom drawings?* Paper presented at the annual meeting of the American Educational Research Association.
- Black, K. (1991). How students see their writing: A visual representation of literacy. *Journal of Reading*, 35, 206–214.
- Branfman, F. (1972). *Voices from the Plain of Jars: Life under an air war*. New York: Harper Colophon.
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the Multitrait-Multimethod Matrix. *Psychological Bulletin*, 56(2), 81–105.
- Clifford, G. J. (1968). *Edward L. Thorndike: The sane positivist*. Middletown, CT: Wesleyan University Press.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20, 37–46.
- Deguchi, M. (1998). *Elementary art: A comparison of Japan and the U.S.* Unpublished manuscript, Boston College, Lynch School of Education, Chestnut Hill, MA.
- Gamradt, J., & Staples, C. (1994). My school and me: Children's drawings in postmodern educational research. *Visual Arts Research*, 20(1), 36–49.
- Goodenough, F. L. (1926). *Measurement of intelligence by drawings*. Chicago: World Book.
- Goodwin, R., & Fahnestock, A. H. (2002). *The day our world changed: Children's art of 9/11*. New York: Harry N. Abrams.
- Golomb, C. (1992). *The child's creation of a pictorial world*. Berkeley: University of California Press.
- Grugeon, E. (1993). Gender implications of children's playground culture. In P. Woods, & M. Hamersley (Eds.), *Gender and ethnicity in schools: Ethnographic accounts* (pp. 11–33). London: Routledge.
- Gulek, C. (1999). *Using multiple means of inquiry to gain insight into classrooms: A multi-trait multi-method approach*. Unpublished doctoral dissertation, Boston College, Chestnut Hill, MA.

- Haney, W. (1977). *A technical history of the National Follow Through Evaluation*. Cambridge, MA: Huron Institute.
- Haney, W., & Madaus, G. F. (1989). Searching for alternatives to standardized tests: Whys, whats, and whithers. *Phi Delta Kappan*, 70, 683–687.
- Heller, S. (2002, September 8). Smoke, grief and dinosaurs [Review of the book *The day our world changed: Children's art of 9/11*]. *New York Times Book Review*, p. 11.
- Houston, J. (Ed.). (1986/1990). *Thesaurus of ERIC descriptors*. Phoenix: Oryx Press.
- Jaeger, R. M. (Ed.). (1997). *Complementary methods for research in education* (2nd ed.). Washington, DC: American Educational Research Association.
- Krueger, R. A. (1994). *Focus groups: A practical guide for applied research* (2nd ed.). Thousand Oaks, CA: Sage.
- Kvalseth, T. O. (1989). Note on Cohen's kappa. *Psychological Reports*, 65, 223–226.
- Lagemann, E. C. (1986). *Contested terrain: A history of education research in the United States, 1890–1990*. Chicago: Spencer Foundation.
- Lifford, J., Byron, B., Eckblad, J., & Ziemann, C. (2000). Reading, responding, reflecting. *English Journal*, 89(4), 46–57.
- Lubin, B., Larsen, R. M., & Matarazzo, J. D. (1984). Patterns of psychological test usage in the United States. *American Psychologist*, 39, 451–454.
- Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (pp. 13–103). New York: Macmillan.
- Mumford, L. (1934). *Technics and civilization*. New York: Harcourt Brace.
- Pressman, J. L., & Wildavsky, A. (1973). *Implementation: How great expectations in Washington are dashed in Oakland; or, Why it's amazing that Federal programs work at all, this being a saga of the Economic Development Administration as told by two sympathetic observers who seek to build morals on a foundation of ruined hopes*. Berkeley: University of California Press.
- Rogers, R. S. (1969). *A consolidated report on the "Draw-A-Classroom" test: A study of the drawing behavior of children in Toronto public schools*. Unpublished manuscript.
- Rosenthal, R., & Jacobson, L. (1968). *Pygmalion in the classroom*. New York: Holt, Rinehart & Winston.
- Rudner, L. M., & Boston, C. (1994). Performance assessment. *ERIC Review*, 3(1), 2–12.
- Russell, M. (1999). *Validity and reliability of information gleaned from student drawings*. Paper presented at the annual meeting of the American Educational Research Association, Montreal.
- Sack, T. (1997). *Using drawings and interviews to promote reflection on student teaching*. Unpublished doctoral dissertation, Boston College, Chestnut Hill, MA.
- Siegel, B. S. (1986). *Love, medicine and miracles*. New York: Harper and Row.
- Steele, C. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist*, 52, 613–629.
- Thomas, G. V., & Jolley, R. P. (1998). Drawing conclusions: A re-examination of empirical and conceptual bases for psychological evaluation of children from their drawings. *British Journal of Clinical Psychology*, 27, 127–139.
- Tovey, R. (1996). Getting kids into the picture: Student drawings help teachers see themselves more clearly. *Harvard Education Letter*, 12(6), 5–6.
- Weber, S., & Mitchell, C. (1995). "That's funny, you don't look like a teacher." Washington, DC: Falmer Press.
- Wheelock, A., Bebell, D., & Haney, W. (2000). *What can student drawings tell us about high-stakes testing in Massachusetts?* Retrieved December 15, 2000 from <http://www.tcrecord.org/Content.asp?ContentID=10634>.
- Wilson, B., & Wilson, M. (1977). Children's story drawings: Reinventing worlds. *School Arts*, 78(8), 6–11.
- Winner, E. (1989). How can Chinese children draw so well? *Journal of Aesthetic Education*, 23(1), 41–63.

The work reported here was supported in part with grants from the New American Schools Corporation, and the Edna McConnell Clark and Spencer foundations. For such support we are extremely grateful, but note that the views expressed here are not necessarily those of anyone other than the authors. We also express our thanks to the students and teachers in schools in Massachusetts, Minnesota, Texas, California, Indiana, and Connecticut with whom we have worked over the past decade or so. Additionally, we thank Chip Morrisson, Cat Xander, Amy Bebell, James Caradonio, Frank Lentvorsky, Jayne Griffin, Carol Pelletier, Catalina Montes, Alyce Anderson, Robert Coles, Larry Cuban, Hayes Mizell, Rosaena Garza, Phyllis Lyons, Les Morse, Toni Cronin, and Steve Delapp for their support of our work with student drawings. Finally, we thank colleagues who have worked with us on one or more of the endeavors reported here: Cengiz Gulek, Anne Wheelock, Ed Fierros, Phyllis Misite, Trudy Sack, Jean Lifford, Mimi Coughlin, Ann Tan, Nicole Malec, Larry Ludlow, Amie Selden, Lisa Jackson, Alka Arora, and Deborah Berger.

This article has been reprinted with permission of the *Harvard Educational Review* (ISSN 0017-8055) for personal use only. Posting on a public website or on a listserv is not allowed. Any other use, print or electronic, will require written permission from the *Review*. You may subscribe to *HER* at www.harvardeducationalreview.org. *HER* is published quarterly by the Harvard Education Publishing Group, 8 Story Street, Cambridge, MA 02138, tel. 617-495-3432. Copyright © by the President and Fellows of Harvard College. All rights reserved.