

Issues in Portfolio Assessment: Assessing Writing Processes From Their Products

Maryl Gearhart and Shelby A. Wolf

*The National Center for Research on Evaluation, Standards, and
Student Testing (CRESST)
University of California, Los Angeles
University of Colorado, Boulder*

We investigated the ways that portfolio evidence of students' competencies with writing processes was created and interpreted in 4 classrooms. Our study was conducted during preliminary classroom trials of California Learning Assessment System portfolios, when teachers and students were challenged with the new task of preparing portfolios that demonstrated students' competency with the "dimensions of learning." Drawing data from teacher and student interviews as well as portfolios, we considered three issues regarding the meaning of portfolio indicators of writing processes: (a) Students' opportunities to learn to use a range of resources, processes, and standards in ways that enhance the effectiveness of their writing; (b) students' opportunities to produce "hard copy" evidence of their uses of processes; and (c) students' capacities to analyze their writing processes. Further research is needed to understand how participants in a large-scale portfolio assessment program develop shared understandings of the ways that evidence of writing processes is considered in the scoring and how the programmatic needs for comparability of evidence can be reconciled with the personal needs of young writers, whose uses of processes will vary with the purposes and contexts of their writing.

Working to piece together the puzzle of writing portfolios, teachers and researchers across the nation are challenged by questions of portfolio design and purpose. This article isolates but a piece of that puzzle: We examine issues surrounding the

assessment of students' competencies with the writing process, focusing on the ways these issues take form in the context of large-scale portfolio assessment. The challenge of interpreting students' competencies with writing processes based on written artifacts is considerable. Purves (1992) questioned

What is it that is assessed? How can we talk about students being better or worse planners, drafters, revisers, or editors? In each case, it seems we have to look beyond the act to the result of the act: the plan, the draft, the revision, or the edited copy. (p. 113)

Purves's words remind us that portfolio artifacts are products of processes, and thus artifacts can never be other than indirect evidence of writing processes that are inherently emergent, cognitive, and interactive.

Indeed, despite general consensus in the literacy community about the writing process—as an iterative, complex experience, rather than a linear sequence of steps that writers must follow to achieve a product (Atwell, 1987; Dyson & Freedman, 1991)—we remain without analogous consensus on methods for portfolio assessment of process. Perhaps reflective of the challenge, process may not be represented in the rubrics or the portfolio specifications developed for large-scale portfolio assessment (e.g., Vermont Department of Education, 1990, 1991). Programs that do incorporate a process component are not yet well researched—for example, the New Standards Project (Murphy, Bergamini, & Rooney, this issue; Simmons & Resnick, 1993; Spalding, 1995). The principal findings available at this time come from studies of the Pittsburgh School District (Arts Propel) portfolio assessment program. Students' "use of processes and strategies for writing" are assessed with a six-level rubric; raters are asked to consider "effective use of prewriting strategies, use of drafts to discover and shape ideas, use of conferencing opportunities to refine writing (peers, adult readers), and effective use of revision (reshaping, refocusing, refining)" (LeMahieu, Eresh, & Wallace, 1992). Technical studies have demonstrated that Pittsburgh raters achieve satisfactory levels of agreement when rating students' uses of processes and strategies (LeMahieu, Gitomer, & Eresh, 1995). However, there exists no coordinated research on Pittsburgh raters' procedures for interpreting process evidence or on teachers' and students' understandings of the ways that raters interpret portfolio evidence.

This article addresses the need for research on the large-scale assessability of writing processes based on portfolio evidence. Drawing on interviews with teachers and students conducted during the field trials of the portfolio program of the California Learning Assessment System (CLAS; Sheingold, Heller, & Paulukonis, 1995), we compare the ways that portfolio evidence of students' competencies with writing processes was created in four classrooms, and we reflect on the ways that different kinds of portfolio evidence of writing processes may be interpreted by raters outside the classroom.

BACKGROUND ON THE CLAS PORTFOLIO ASSESSMENT PROGRAM

In collaboration with the California Department of Education, the Center for Performance Assessment of the Educational Testing Service (ETS) was engaged from 1992 to 1994 in the design of a standards-based portfolio component for CLAS. Working with educators across the state, ETS was developing a portfolio assessment system that could build on and support improved classroom practice, while also providing trustworthy information about student performance. The approach that evolved focused on students' demonstrating performance with respect to "dimensions of learning" aligned with the instructional goals of the California state frameworks.

In language arts, the rubric was designed to capture competency with two broad dimensions: (a) composing and expressing ideas and (b) constructing meaning. Students' competencies with process—our focus here—were among the list of bulleted considerations for a rater's decision on composing and expressing ideas:

[Students] draw on *various resources* [italics added], including people, print and non-print materials, technology, and self-evaluation to help them develop, revise and present written and oral communication. They engage in *processes* [italics added], from planning to publishing and presenting; when appropriate, they do substantial and thoughtful revision leading to polished products. Through editing, they show command of sentence structure and conventions appropriate to audience and purpose. (Thomas et al., 1995, p. 11)

Raters in CLAS trial scoring sessions were asked to consider evidence of the writing process, and participating teachers were asked to help students provide that evidence.

GOALS OF OUR STUDY

Our study was designed to reveal issues surrounding the creation, selection, and interpretation of process evidence in CLAS portfolios. Three questions guided our inquiry:

1. *Opportunities to learn*: How did instructional and assessment practices support students' uses of resources and a range of writing processes in ways that enhanced the effectiveness of their writing? We gathered evidence from teacher interviews and from portfolios to understand the ways that classroom contexts supported students' engagement with the processes of writing.

2. *Opportunities to produce "hard copy" evidence of learning*: What kinds of artifacts emerged from classroom practices, and what potential did these artifacts have as evidence of students' competence with various writing processes? We

compared teachers' reported opportunities to learn (based on the interviews) with the hard copy artifacts in the portfolios.

3. *Students' capacities to analyze their writing processes*: How did students explain their writing processes, either in writing or in discussions with us? We asked students to discuss selected pieces of portfolio evidence, and we examined relations among their interview explanations, written reflections, and the process artifacts contained in the portfolios.

From these findings, we drew implications regarding the capacity of a large-scale portfolio assessment program to provide meaningful indicators of students' competencies with writing processes. Because we collected our data during field trials of a program that was never implemented, we close with regrets that the demise of CLAS made it impossible for us to follow these and other classrooms forward in time.¹ Research is needed to understand (a) how teachers and students participating in a large-scale portfolio assessment program can develop shared understandings of the ways that evidence of writing processes is considered in the scoring and (b) how the programmatic needs for comparability of evidence can be reconciled with the personal needs of young writers, whose uses of processes will vary with the purposes and contexts of their writing.

METHODS

The four study classrooms (Grades 2, 4, 7, and 8) spanned urban, rural, and suburban settings. Six target students in each classroom were selected by each teacher to represent the diversity of ethnicity, gender, and language arts competencies (two high, two medium, and two low) at each school site. (See Table 1.) Attending from one to three meetings over a span of 5 months, the four teachers had contributed to the formative design of the CLAS-ETS portfolio assessment system—the development of the dimensions of learning and the assessment guides, recommendations for classroom implementation, and trial portfolio scoring. Two teachers had sent their students' completed CLAS language arts portfolios to the trial scoring session just prior to our visit, and the remaining two teachers were engaged with their students in preparing portfolios consistent with the CLAS model at the time of our visit.

Our teacher and student interviews were refined in pilot interviews with two elementary teachers and six students; in addition, each of the four participating CLAS teachers suggested minor revisions of the student interview appropriate for their students. The interviews focused on students' writing (rather than all components of the English and language arts dimensions of learning) and contained language from the dimensions of learning; Table 2 contains the questions pertaining to writing processes of relevance to the findings reported here. The teacher

¹In September 1994, California Governor Pete Wilson vetoed a bill reauthorizing all components of the CLAS Program

TABLE 1
 Teachers and Students at the Four Study Sites: Predominant SES and Ethnicity

<i>Teacher</i>	<i>Setting</i>	<i>Students</i>			<i>Note</i>
		<i>Grade Level</i>	<i>SES</i>	<i>Ethnicity</i>	
Aimes	Rural-suburban	2	Working class	Predominantly Anglo	
Bentley	Suburban	4	Middle class	Predominantly Anglo	
Cris	Urban	7-8	High poverty to middle class	Ethnically diverse	Participant in trial of New Standards portfolio assessment
Donner	Rural	8	Working class	Predominantly Anglo	

Note SES = socioeconomic status.

interview was supported by relevant artifacts (assignment sheets, editing guides, rubrics, sample student portfolios). The student interview was supported by writing from the student's own CLAS portfolio. Interviews were transcribed from audio-tape, and copies were made of each student's portfolio.

Our analyses of interviews and portfolio contents were designed to be generative to fit the context of our study. Because the CLAS portfolio program was in its formative phase and the four classrooms varied markedly in characteristics, we viewed our data as a resource for producing frameworks for investigating what teachers and students understood of the CLAS dimensions of learning and the ways that their understandings mediated curriculum and portfolio choices.

Following Erickson (1986), our analyses represented our search for *key linkages* in the entire data set, including interview responses from teachers and students, students' written reflections on their writing processes, and process artifacts. Erickson viewed the goal of qualitative research as the construction of a complex of relations among general assertions and subassertions based on evidence from

TABLE 2
Interview Questions

<i>Teacher Interview Questions</i>	<i>Student Interview Questions</i>
<p>What resources do the students draw from to inform ___ (each of the genres, purposes mentioned)? How do they learn to take the things they know and read and connect these with their own writing? How are students learning to use resources to develop, refine, and present ideas? How do they show this growth in their portfolios?</p>	<p>Where did you get the ideas for this—from your teacher, from other kids, or other books?</p>
<p>How are students learning to “use a range of processes from planning to revising, editing, and presenting?” Tell me about the ways you implement a writing process approach</p>	<p>Is there a piece where you did a lot of revision or really worked hard to change it? For this piece, once you had an idea, what did you do? Once you'd written something, did you revise? How?</p>
<p>How are students learning to “ .[apply] explicit standards for judging the quality of their own and others' work?” How do they show this growth in their portfolios? What are the sources for students' assessments of their writing? From classroom standards? How are students informed of these standards? Are some required and others negotiated? Do students gain understandings of their writing from peers? How? From parents? How?</p>	<p>When you revise, do you use classroom guidelines or the responses of your teacher or fellow students to help you? How? Do you ever help other students revise—read or listen to their work and give them ideas? What kind of advice do you give? Can you give me an example?</p>

multiple and varying data sources. As researchers scan their data, they develop an analysis of the patterns of evidence—connecting, for example, fieldnotes to site documents to audiotaped transcripts, or, as in our case, connecting what teachers and children said with what they wrote.

Figure 1 illustrates our methods. Based on our analysis of a pattern of evidence, we claim that, although all four teachers emphasized substantive revision in the interviews, the student interviews and the portfolios revealed that students' under-

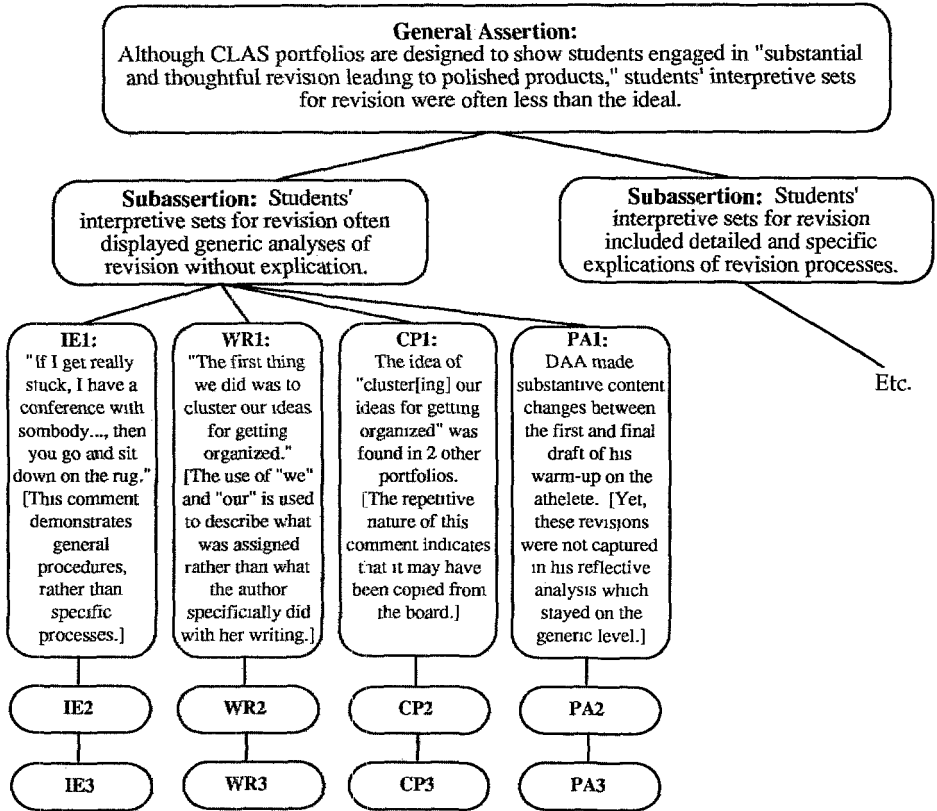


FIGURE 1 Key linkage example. CLAS = California Learning Assessment System, IE = Interview Explanation (students' and teachers' interview discussion with researchers); WR = Written Reflection (forms or letters to teachers explaining revision), CP = Comparing Portfolios (comparing reflection comments among students in the same class); PA = Process Artifacts (comparing drafts of specific pieces over time within one student's portfolio). From Erickson, F (1986). "Qualitative Methods in Research on Teaching." Adapted with permission of Macmillan Library Reference USA, a Simon & Schuster Macmillan Company, from M. C. Wittrock (Ed.), HANDBOOK OF RESEARCH ON TEACHING, pp. 119-161. Copyright © 1986 by the American Educational Research Association.

standings of revision often differed from the ideal. A subassertion—one component of the larger claim—was that students’ analyses of their revisions were often generic in nature. We support this subassertion with several categories of evidence: In interviews (IE1) as well as in written reflections (WR1), students often reported how “we” revise without showing how or why they personally revised particular pieces; reflective writing often contained much the same content across portfolios (CP1); students often demonstrated substantial revision over several drafts, and yet their written reflection on the changes failed to capture these changes (PA1). This example does not represent the full range of data and analytic methods but provides the reader with a taste of the diversity of our resources and our approach to discovering linkages across resources.

In our search for patterns as well as discrepant cases, Shelby Wolf assumed greater responsibility for analysis of opportunity to learn, MaryI Gearhart had greater responsibility for students’ opportunities to produce evidence of learning, and both of us shared analysis of students’ capacities to analyze their writing processes. However, each of us examined all evidence and challenged one another’s interpretations as we compiled our findings into key linkages.

FINDINGS

In the sections that follow, we organized our findings around our three core questions. Evidence of opportunity to learn is drawn from teacher interviews and portfolio materials. Evidence of opportunity to produce hard evidence is based on our comparisons of teachers’ reported practices with the artifacts we found in the portfolios. Evidence of students’ capacities to analyze their writing processes is based on analyses of the student interviews and portfolios.

A companion study based on additional data collected in the same four CLAS classrooms set the context for the analyses we report here (Gearhart & Wolf, 1994; Wolf, Davinroy, & Gearhart, 1997). Focusing on the portfolio assessment of students’ understandings of writing purposes, genres, and audiences, we reported that the four participating teachers varied in their understandings of the CLAS dimensions in ways that helped explain the opportunities to learn that they provided their students as well as the opportunities their students had to produce hard evidence for their portfolios. We interpreted patterns in the teachers’ assignments in the context of tensions between the “romantic” and “classical” schools of composition theorists (Hairston, 1986). In the romantic view, students must write from their own questions and emotions to make their own meaning in the world; in the classical view, students are taught to analyze many kinds of writing as a grounding for their efforts to extend their range and flexibility as writers. We questioned our participating teachers about the extent to which their interpretations

of purpose, genre, or audience reflected a relatively balanced perspective between the romantic and classical schools (the balanced perspective of the CLAS dimensions of learning), or whether they leaned more heavily toward either pole. We found that the curriculum of three teachers was somewhat more classical: Ms. Bentley, Ms. Cris, and Ms. Donner developed assignments as genre studies and assigned specific authentic or imaginary audiences, and their students' portfolios were built on the teachers' assigned tables of contents. In contrast, the curriculum of a fourth teacher, Ms. Aimes, was more romantic: She placed emphasis on writing from personal experience, and her students' portfolios were a matter of student choice.

We recognize that teachers' inclinations toward either the romantic or classical dimensions were associated with the grade level they taught. Ms. Aimes was a primary teacher who emphasized the need for her students to write from personal experience and deemphasized genre studies, assignment-specific rubrics, and multiple drafts. The upper grade teachers placed a strong focus on the forms and functions of particular kinds of writing, structured means of reflective assessment, and the multiple iterations necessary for young writers to accomplish effective pieces. We represent this association between philosophy and grade level in our small sample as a weave of grade level and orientation (see Figure 2): The orientations of the teachers in this study were neither solidly romantic nor classical, but each made her way toward the balanced vision that CLAS offered from the perspectives of her own grade level and prior philosophies. Our weave thus represents the possibility of association between grade level and philosophy. Although one might assume that Ms. Aimes held her views because she taught in the primary grades, often romantic in both philosophy of development and orien-

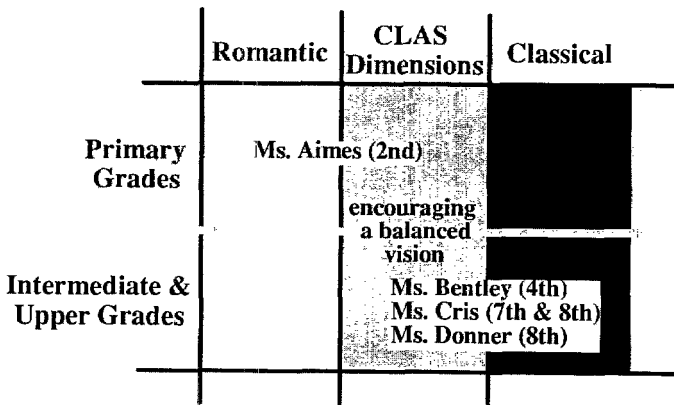


FIGURE 2 The need for a balanced vision. CLAS = California Learning Assessment System.

tation toward instruction (McGill-Franzen, 1993), her emphasis on personal purposefulness is not limited only to teachers in the early elementary grades. Indeed, we have previously discussed a case of an upper grade teacher who was just as committed to a view of writing as a forum for personal purposes and just as worried that explicit genre instruction might dampen the spirits of his student writers (Gearhart & Wolf, 1994; Wolf & Gearhart, 1997).

In this study, we expected the teachers' views of purpose, just summarized, to be related to their views of the process of writing, for writing processes support the effectiveness of a piece for its intended purpose. The concept that purpose and process must coevolve for a piece to be effective is a very complex frame for viewing process, but it captures how more accomplished writers work. For example, writers may begin with issues and audiences they want to address within particular genres and forms. As they write, they are constantly calibrating their words to ensure that they are achieving their intended purpose, and they ask for advice from others (e.g., friends, editors, reviewers) who challenge them to refine their thinking even more. Thus, a writer's—or a writing teacher's—view of purpose will be reflected in the ways that he or she uses particular processes in the development and refinement of any piece of writing.

Opportunities to Learn Writing Processes and Opportunities to Produce Hard Evidence of Learning

Table 3 outlines our findings regarding students' opportunities to learn uses of resources and a range of processes and standards as guides for the analysis of one's writing. The table is organized to reveal the sources of our evidence: *Portfolio-documented* opportunities to learn were in evidence in the portfolio and were also discussed by the teacher in our interviews; *interview-documented only* opportunities to learn were documented only in the interviews. The purpose of this distinction is to highlight one of our strategies for identifying missing evidence of learning in the portfolios. Later, we summarize our findings regarding opportunity to learn and then utilize the distinction between portfolio-documented and interview-documented only evidence to consider students' opportunities to produce portfolio artifacts.

Opportunities to learn. From the interviews, we learned that each of the teachers emphasized the human resources of self, by encouraging children to generate their own ideas for writing, and of teacher, through their own instructional leads and written comments. However, all teachers placed the most emphasis on the influence of peers. As Ms. Bentley commented, "They're a lot more interested in ... doing a better job if they know they're going to be evaluated by their peers,"

and Ms. Donner said, “They’re most interested in their peers, so they really rely on what their peers think and what their peers have read, and it’s very effective.”

The amount and purpose of peer conferencing reflected either more classical or more romantic orientations. Ms. Aimes encouraged any student who was uncertain about his or her writing to consult another student, although she left open whom to consult and for what purpose: “Once in a while I’ll have a kid that’ll say ‘I’m stuck. I need help.’ Most of the time they figure it out in peer conferences.” With a more classical view, the other teachers facilitated specific and formal strategies for peer response for studies of genre or the CLAS dimensions of learning: For example, Ms. Bentley guided her class in the creation of rubrics for peer evaluation when they were studying the persuasive letter, Ms. Chris expected peer evaluation of the cover letter to their portfolio, and Ms. Donner required peer response when students were writing to demonstrate the ability to take on varying perspectives.

Students’ uses of prose and poetry models also reflected either more classical or more romantic orientations. Teachers with a more classical view often taught genres by reading and discussing exemplars. As Ms. Cris explained, “I ask the kids to look at the way authors use words in context, or the way they write dialogue.” In contrast, Ms. Aimes, the teacher whose curriculum we regarded as more romantic, encouraged the use of wordless picture books and pattern books as “real comfortable support” for personal writing, support that took form as individual transformations of established patterns. For example, students were asked to translate *The Important Book* (Brown, 1949)—a set of ideas that children generally find important—into images of what was important to them personally. Students’ uses of traditional writing resources were in evidence in all classrooms: Children wrote definitions for words, edited their spelling, and altered their word choices with the help of a dictionary or thesaurus.

The opportunities to learn to use a range of processes provided in the four classrooms were quite extensive, and there were some strong commonalities across classrooms. From both the interviews and the portfolios, we learned that students were engaged in planning (e.g., webs, matrices, warm-ups), drafting (various genres written in sequential drafts), using the visual arts (e.g., sketches, illustrations) to accompany and extend the written words, revising (attention to content changes, organization, audience, genre), and editing (attention to spelling and mechanics). To help students deepen their understandings of the interplay between purpose and process, all four teachers provided students with standards to guide the writing process—guidelines for planning, examples of good work or work that needs improvement, rubrics for structuring reflection on the qualities of good writing, checklists, and other artifacts to support final editing (word changes, polishing of mechanics). The teachers developed and used standards for processes in ways that reflected both their curriculum and their beliefs about students’ capacities to use standards to guide planning, composing, revising, and editing.

TABLE 3
Evidence of Students' Opportunities to Learn

<i>Opportunities to Learn</i>	<i>Romantic Primary Grade Teacher</i>		<i>Classical Intermediate and Upper Grade Teachers</i>	
	<i>Portfolio-Documented Evidence^a</i>	<i>Interview Evidence Only^b</i>	<i>Portfolio-Documented Evidence^a</i>	<i>Interview Evidence Only^b</i>
Use of resources	<p><i>Self:</i> students' own ideas</p> <p><i>Teacher:</i> teachers' instruction and written comments</p> <p><i>Structured models from prose and poetry:</i> patterns in student writing in match well-known patterns in professional writing</p> <p><i>Dictionary and spellers:</i> definitions written, spelling edited</p>	<p><i>Discussion of specific books and patterns:</i> no audiotapes of book discussions or literary response journal entries; no attribution to professional sources</p> <p><i>Peers as informal resources:</i> no written peer commentary</p>	<p><i>Self:</i> students' own ideas</p> <p><i>Teacher:</i> teachers' instruction and written comments</p> <p><i>Peers as formal resources:</i> written peer commentary complimenting and criticizing student work</p> <p><i>Generic models from prose and poetry:</i> attribution to professional writing; rewriting well-known scenes</p> <p><i>Dictionary and thesaurus:</i> definitions written; word choice altered</p>	<p><i>Discussion of texts as exemplars of particular genres:</i> no audiotapes of class discussion, no double entry diaries or literary response logs that analyze the craft of professional writers</p>

Use of processes
and standards

Planning brainstorming webs
with small pictures

Drafting: stories, poems,
personal narratives, usually
limited to one draft and the
final copy

Using the visual arts: sketches
and crayon drawings that
accompany and extend the
written text

Revising minor attention to
content changes

Editing. checklists for student
editing with focus on
mechanics; teacher editing of
mechanics on first drafts

*Reflection on growth as a
writer*. students' letters to
teacher about contents of the
portfolio listing favorite
pieces and discussing growth
in generalized ways

Peer conferencing no written
peer feedback; no
self-reflective writing to
indicate that peer criticism
was taken into account

Group share. no audio- or
videotaped presentations, no
"Author's Day" invitations;
no overheads summarizing
class discussion of a
particular piece

*Teacher conference on
revision*: no written criticism
of students' writing other
than general comments like
"Good job!" or "Try again!"

Peer conference on revision.
no written commentary from
peers

Planning topical maps,
matrices, warm ups

Drafting. stories, poems,
personal narratives, and
exposition usually in
multiple drafts in formal
sequence

Using the visual arts: drawings
and three dimensional
artwork to accompany
written text

Revising: some self-, peer, and
teacher attention to content
changes, often guided by
assignment-specific analytic
rubrics

Editing: self-, peer, parent, and
teacher criticism of spelling
and mechanics, often guided
by editing checklists

*Reflection on growth as a
writer*: student writing that
sets goals for addressing
criticism, students'
explanations of portfolio
selection choices
highlighting favorite pieces,
a few portfolio entries
exploring process

Oral peer reflection no
audiotapes of "extended
conversations" analyzing
students' work

Analyzing writing: no
overheads summarizing class
discussions analyzing
student or professional
writing

*Understanding differences
between revision and
editing*. little substantive
content revision beyond
word substitution (although
teachers said students
understood the differences
well)

^aPortfolio artifacts in evidence. ^bNo portfolio artifacts that document teachers' reports

Thus, to organize their prewriting and “brainstorming,” the students in Ms. Aimes’s more romantic classroom were encouraged to draft webs and suns containing a large category in a center circle and then lines leading outward indicating subcategories—for example, *birthdays* in the center circle with lines for *pizza*, *friends*, and other subtopics accompanied by small illustrations.

We talk about brainstorming ... that you just get everything out that you know, and it’s a way to start thinking.... I’m not sure at this point yet how many of them actually then look at [their prewriting] when they go to [write], although I make sure that they have all the papers back when they’re writing their rough draft.

Ms. Aimes encouraged webs and suns but did not require their use as resources for the writing; her standard was, “It helps to think ahead.”

The more classical teachers were more likely to engage students in particular writing strategies—a required series of phases or number of drafts, peer response to a draft guided by a prepared set of questions, peer and self-assessment using a rubric—as well as reflection on the writing process. We found a greater range of artifacts in the portfolios, reflecting a more formal curriculum for process phases and strategies, as well as greater consistency in process documentation. At the elementary level, Ms. Bentley designed detailed guidelines for most assignments. Thus, she viewed the objectives for a persuasive letter assignment as the “develop[ment of] an understanding of writing persuasive letters using the writing process.” To carry out this goal, she gave students directions to follow a specific set of steps, including “Note-taking (brainstorming), Published forms for planning, Parent editing of rough draft, Final draft and evaluation.” She and her students also designed an end-of-assignment rubric used throughout the year that placed emphasis on both content (“Creative, Descriptive, Use of Prior Knowledge”) and mechanics (“Neat/Organized, Followed Directions, Spelling, Punctuation/Sentences”). At the conclusion of many assignments, students were asked to compare their scores with those of their peers and Ms. Bentley (“How do you compare what you think you need to improve on? Or what are you proud of?”).

The middle school teachers fostered an understanding of writing as an ongoing process, and therefore multiple drafts and content revisions were commonly required. As Ms. Cris explained, “At the beginning of the year, they’ll hear me say really ridiculous things like, ‘You’ll never write anything once.... You’re never done.’” Ms. Cris devised a variety of strategies to support students’ approaches to revision. For example, she “encourage[d] them to draw lines through their errors and not erase, because I often like to see what it was they chose to take out.... [S]ome of them ... think a sign of growth is that they erase less on the next draft.” To encourage student reflection on the writing process, she trialed a method from the New Standards Project, an opportunity to choose a “Process Entry” and reflect

in writing on the original idea (“How did you get started?”), emergent composing (“How did you go about the writing of it? How did your ideas change?”), challenges (“What were the hardest decisions?”), major revisions (“What were the biggest changes?”), and evaluation of the finished piece (“What do you think about how the piece turned out?”).

The very idea of revision was a difficult concept for Ms. Donner’s students: “Because I find at this age, ... they just struggle against revision. They like it the way it is the first time.” Concerned, she challenged her students with the standard of “showing” writing for every assignment—writing that utilized the five senses in specific ways.

I said, “I want you to be different; I want you to really focus on making changes that are better, instead of just changes.” ... What they had to do is say exactly what they changed. And some of them said, “I changed the word ‘frog’ to ‘toad’ in paragraph two, and I spelled ‘their’ correctly.” And those were okay, but I told them that I was looking for more.... If they didn’t do much “showing” writing, or didn’t take advantage of what other students said, then I didn’t award anything.

Ms. Donner hoped that her press for showing writing would move her students beyond the kinds of local changes that were typical in CLAS portfolios she had reviewed from other classrooms.

Opportunities to produce portfolio evidence of processes. Comparisons of the opportunities teachers reported in their interviews with the evidence of processes contained in the portfolios revealed gaps between what students were learning to do with text and the available hard copy evidence of their learning. Consider documentation of students’ uses of textual resources. Although the more classical teachers described to us the texts they used as exemplars of particular genres and the more romantic teacher explained her use of specific books, none of the teachers placed emphasis on students’ documentation of these resources, and thus, in the portfolios, texts were unpredictably named and their uses rarely analyzed in students’ writing. A child’s use of text can range from a slim fragment of character or setting description to a heavy borrowing of entire chunks of text (Wolf & Heath, in press). With little or no documentation of the child’s uses of texts, raters’ evaluations would thus depend solely on their knowledge of literature.

A pattern emerged of imbalance of opportunities to select evidence across the range of writing processes. Evidence of editing was far more abundant and far easier to interpret than evidence of the kinds of complex, content-based planning and revision processes emphasized in the CLAS dimensions. The forms teachers borrowed or created to support writing processes—and consequently, the artifacts

that made their way into the portfolios—often focused on surface editing and looked more like the artifacts of past skills-based curricula than the balanced perspective advocated by CLAS–ETS. Thus, use of texts as resources and models and revision processes that teachers reported supporting and that students reported utilizing were underrepresented in the portfolios. The documentation that did exist was often teacher structured in problematic ways. For example, in analyzing his or her planning or content revisions, a student might parrot a closed prompt (“Yes, I did include prior knowledge”) or provide an undeveloped response to an open prompt (“I will write more next time”) that left us wondering about the student’s conception of a substantive response. Thus, the portfolio evidence we found was often a distortion of the richer opportunities for writing processes provided in the classroom.

We recognize that the CLAS program was new and the teachers’ experiences with the dimensions were brief, but inexperience was only one source for this imbalance of evidence. Another source was the teachers’ philosophic commitment to particular roles for teachers and students in the context of composing. The teachers had different views of the teacher’s role in promoting growth in authorship, and their philosophies motivated their choices to impose (or not) or request (or not) particular forms of documentation from students. Indeed, what was absent in the portfolios was sometimes the by-product of well-motivated practices in the classroom, and thus, there were conflicts between the needs for portfolio evidence and pedagogical practices.

Consider the more romantic classroom. Ms. Aimes provided a specific support structure for writing (e.g., webs), but whether a child used that support later was a matter of personal artistic preference. Her pedagogy—thoughtfully grounded in her framework for primary-level writing—had consequences for the evidence her students had available at the time of portfolio construction. A student might choose to include a web or sun that he or she had never actually used; a student might choose to exclude a web or sun that he or she had eventually decided was irrelevant. A dilemma for portfolio assessment emerges if the score is intended to reflect an individual student’s competence with the writing process: How could a rater discern how a child from Ms. Aimes’s room used (or decided not to use) a web or sun?

Ms. Aimes also provided students with opportunities to solicit or provide peer input during the planning or composing of a piece, consistent with the value that the CLAS dimensions of learning placed on students’ understandings of the roles of peers as readers. Yet, these opportunities were never documented. Ms. Aimes had not devised a way to provide evidence of the complex processes of working through a piece with a peer, discussing possibilities, and setting down next steps; as a result, she did not have much knowledge of how these conferences were working: “Honestly I have not sat down and listened to them. I don’t know.” Thus, the portfolios contained no artifacts reflecting the processes and outcomes of peer conferencing and “group share”—a whole-class meeting in which one student reads

his or her story and the other students “tell back” what the author has read, make critical commentary, “and then offer suggestions or ask questions.” However, Ms. Aimes did require that students use standard artifacts for editing: “They know that there’s a difference between what they do in group share, when they’re asking for advice and feedback, and what they do when they have to fill out this sheet.” The edit sheet centered on punctuation, spelling, and the kinds of grammatical errors that are often targeted when a paper is read aloud.

In the more classical rooms, there was greater likelihood of portfolio documentation at all phases of the writing process. However, the existence of such “managed portfolios” raised countervailing concerns that a teacher’s imposed procedures for portfolio documentation may obscure rather than reveal students’ use of processes and their understandings of process.

Consider the documentation we found in Ms. Bentley’s portfolios. With the support of either a published or a class-constructed rubric, the fourth graders in Ms. Bentley’s class were provided with opportunities to reflect on the processes of creating almost every writing assignment. Midwriting self-evaluations were usually supported with published rubrics, most of which focused more on mechanics than content. In the “self-evaluation master” on “Writing a Persuasive Letter,” the majority of the 17 questions dealt with issues of form, spelling, and neatness: “Did I use the correct form for a letter? Did I spell all words correctly? Did I copy my letter neatly and correctly?” The four that dealt with content were *yes-no* choices that remained on the surface of writing—on the outward features of particular genres, rather than on the more complex uses of language to achieve these features: “Does the body of my letter begin with a topic sentence? Does my topic sentence give my opinion? Are all the reasons for my opinion clear? Do I have a strong last sentence?” Students’ responses made evident the limitations of such *yes-no* questions as prompts for process assessment: Even the “best” writers in the class responded by parroting back the questions, “Yes, there is a strong last sentence” (BSL, student writing, n.d.), providing no portfolio evidence of what was understood.

Ms. Bentley expected her students to carry forward to the composing of the next assignment what they had learned from evaluative feedback at the completion of the prior assignment, scaffolding transfer by encouraging students to consider similar issues in each evaluation. However, repetition of these evaluation components across assignments appeared to encourage students’ use of language directly from the rubric, to the neglect of an analysis of the ways that the rubric applied to the processes of composing a specific piece of writing. For example, one student filled out a “My Portfolio and Me” paper, saying that she chose her miner’s letter for a “Showcase Piece” because “(a) It is creative, descriptive; (b) it has use of prior knowledge; and (c) I followed directions” (BMK, student writing, n.d.). Thus, BMK’s written reflection centered more on imitating set goals (the class rubric) than establishing personal understandings of process or quality.

In the middle school grades, Ms. Cris was struggling to find ways to encourage reflective processes without reducing processes to rigid routines or requiring burdensome written analyses.

A lot of kids complain . . . , “When you want us to reflect, it’s always in writing, always in writing, always in writing.” And the thought occurred to me that if you have a student who does not write well, who does not like to write, and then you insult him after he’s created the piece to write a reflection—how ridiculous. So I decided . . . [to] restructure our whole school portfolio process to do what we call student reflections groups, where they come together with a set of three questions, which they design themselves, to present to their group and get feedback in an oral fashion, and they write down the comments that are made about how to improve their work or revise it.

Ms. Cris’s decision, designed to benefit students, had an unintended consequence for external portfolio assessment: The opportunities that Ms. Cris provided for student reflection were impossible for us to track in the portfolios. The conversations of the reflections groups were captured only in cryptic notes: “Do you think I met up to the expectations [of the assignment]?” “You’re kinda in there. Put more into it and you’ll be there.” “To me, I think you missed meeting just barely. Improve a few mistakes and you’ll meet.” In this case, a teacher’s thoughtfully motivated classroom practices limited the evidence of processes available for students’ portfolios.

Summary: opportunities to learn processes and opportunities to produce portfolio evidence. All four teachers provided students with opportunities to learn and to use a variety of resources, processes, and standards when composing their work, but they differed in the writing strategies and standards taught as well as in their approaches to supporting the production of portfolio evidence. Differences could be explained in part by teachers’ philosophic commitments to more romantic or more classical treatments of the writing process, which were often integrated with their goals for the grade level they taught. The more romantic (primary) teacher assured us that rich and varied activities supported her students’ writing (e.g., literary analysis, peer conferences) and were a part of her daily classroom life. Yet, she felt that the documentation of such a classroom life was either beyond the capacity of her children or beyond her own ability to track. On the other hand, the more classical teachers were likely to impose formal and documented requirements for writing processes; however, the middle school teachers then struggled to find a balance between the potential benefits of imposed requirements and the needs for young writers at times to write without documenting

every phase. There were also patterns across all classrooms: Portfolios were more likely to contain documentation of editing than planning or content revisions. In addition, students' written reflections as well as peer responses were likely to be vague or unexplicated repetitions of classroom standards ("Yes, I used ____."). The latter finding raised questions about students' understandings of reflection and analysis as a genre and of the ways that reflective writing can provide evidence of processing, issues that we turn to in the next section.

Students' Capacities to Analyze Their Writing Processes

Over the course of a year, the students we studied prepared evidence for their portfolios. In this section, we consider students' understandings of the ways that portfolio evidence reveals what students know and can do with text. We asked: How do students analyze evidence of their writing processes?

As depicted in Figure 3, in most models of portfolio assessment, evidence of writing processes consists of dated artifacts linked to the various phases of the writing process (e.g., planning webs or suns or outlines, dated drafts showing changes and editing) and explanations of revisions (e.g., written reflections or notes from oral discussions); a rater's task is to make inferences about students' competencies based on relations among all these sources. In our interviews, we did not

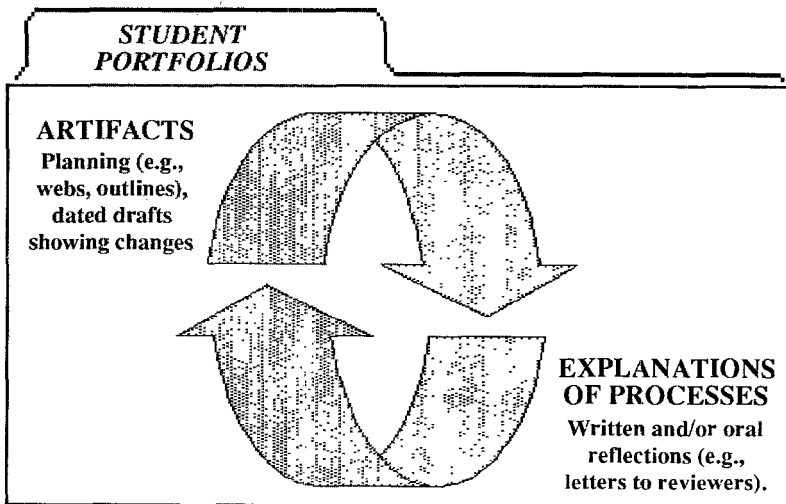


FIGURE 3 Portfolio evidence of processes produced by the student and interpreted by the rater.

TABLE 4
Students' Interpretive Sets for Revision

<i>Interpretive Set</i>	<i>Characteristics</i>
Little analysis of revision in terms of content and purpose and little analysis of readers	Focus on external consequences (grades, praise, global criticism, friendship) Focus on editing of mechanics and on neatness Little analysis of readers as helpful resources Few or no content changes in successive drafts Absence of technical language for any content changes in specific pieces of writing
Vague and inexplicit analyses of revision	Greater focus on the number of drafts than the content changes therein Little analysis of readers as helpful resources Few or no content changes in successive drafts Absence of technical language for any content changes in specific pieces of writing
Generic analyses of revision without explication	Reference to the use of peer response without explication, suggesting the provision or use of peer response was more or less an exercise Few content changes in successive drafts Use of technical language when explaining revisions, but not linked to changes made in the specific piece being discussed
Explications of revision processes	Use of technical language to explain methods, concepts, or standards used to revise and refine the content of a specific piece of writing Evidence of content changes in successive drafts

ask students to consider all evidence but instead focused our interviews on dated drafts of specific pieces, considering processes of content-based revision that can improve the effectiveness of a piece for its intended purpose.² We then analyzed relations between students' discussions of their draft revisions and the process artifacts related to those pieces in the portfolios.

The primary outcome of our analyses is a scheme for characterizing the prevailing patterns of explanations that students provided either in their written reflections or in their interview responses (Table 4). Reflecting the ways that students talked or wrote about revision of particular portfolio pieces, these four categories organize the findings we report, as well as our reflections on the "assessability" of written explanations of process.

²The quality of the portfolio evidence challenged the conduct of our interviews as well as our analyses. As we have reported, the portfolios produced during the initial CLAS field trials were often missing trace artifacts, particularly those linked to content revisions, and students' reflective writing about process was often cryptic or repetitive.

Little analysis of revision in terms of content and purpose and little analysis of readers as resources. Some students talked to us or wrote about revisions of portfolio pieces with little reference to the content of the writing or to the ways that readers' responses could be resources for content revisions. The lengthier or more forceful the response, the more concerned we were about the student's capacity to analyze the content of his or her revisions.

Some of the responses that fit the pattern of "little analysis" focused on external consequences of grades, global reactions, or friendship. Consider the worries of CFR and CPB about grades. CFR explained to us, "If you revise it, you'll get a better grade," a view he presented within his portfolio as well:

I wrote these piece because I wanted to get a good grade and that was the assignment. My language teacher and some of me pears. The friends liked it and my teacher like it. I got a good grade and she told me some ways that I can try it and meke it better. (CFR, student writing, December 3, 1993)

Like CFR, CPB viewed "getting the grade" as the goal: He considered revision as a punishment for a poor first draft ("I really follow her strict directions because I thought she would be mad at me if I didn't. I thought I'd have to do it over or stuff"). Demonstrating little evidence that he pursued his revisions in ways that were grounded deeply in the purpose of his piece, he commented that revisions that did not result in a good grade could be most frustrating:

I revised [my first book report] because I didn't follow directions very well and I still ended up with a 1.5 'cause I didn't read the directions the first time and I wasn't getting it so I kept going and going and I didn't really get a good grade.

Other students discussed the generic reactions of peers or teachers without considering the specific ways that their responses provided useful perspectives on specific drafts. ASJ viewed peers as "people who tell me what they like: ... They say, 'The illustrations are really good and I like the part about' whatever.'" DHH was more concerned about teachers as "people who tell me what they don't like: ... When I showed [this piece] to her she's, like, 'This is way too long, You're supposed to ask me first'. So she started x-ing out the things that I wouldn't need in my final draft." Although we noted that Ms. Donner's response served to transform DHH's first draft from a series of *tell* and *said* to dialogue, DHH interpreted her response more as a personal affront.

Other responses that fit the pattern of little analysis focused on editing of mechanics and on neatness of the work, even when students were free to assess their own or their peers' writing in more substantive ways. This focus on surface-level editing was quite common. For example, one student who consistently

self-evaluated her work in terms of neatness received virtually the same evaluation from peers on one assignment: Three out of four peers commented that her work was “a little messy”; displeased with her final grade, she lamented, “I did not know that we were going to pass it around. I would have done it neater!” (BMK, student writing, December 10, 1993). As another example, when BSL completed an open-ended prompt on her “Favorite Showcase Piece” form, she wrote, “I like it because *My writing is neat, discriptive, and well graded*” (BSL, student writing, November 29, 1993).

Vague and inexplicit analyses of revision. There were occasions when students wrote about or discussed revisions of a piece at length, yet, despite their investment in their piece, it was difficult to glean what they understood about the ways they had revised it. Absent was the use of technical language for characterizing the processes of revision, as well as any analysis that made explicit reference to sections of the text. Such vague explanations raised dilemmas for us when the revisions themselves were more substantial than the student’s explanation of them.

CFR, for example, focused more on the number of times he attempted to revise the organization of his community service piece than he did on the content changes therein; missing technical language for describing his content revisions, CFR’s explanation did not convey deep understanding of the ways that pieces can be substantially reorganized:

I picked this report [for my portfolio] because I knew this was the best one I did all year ... it’s good because I took my time and wrote it and then I recopied like three times before I put it in here.... All the paragraphs had different beginnings.... This is my first one. Second, third, and then I had wrote it different and put the paragraphs in a different way.... Because as I looked through all this right here, it didn’t sound as good as this one did right here. Then I had some of my friends read it over and the next day I wrote this which was better than all these.... Because I know that if you revise it, you’ll get a better grade.

His written reflection was no more revealing of his content revisions:

Well at first I didn’t write what I wanted to so I messed up a couple of times. I learned that I need community [service] no[t] just for school but many other reasons.... My biggest changes was I had to move the paragraphs around so it would look right. (CFR, student writing, n.d.)

However, we found greater evidence of competency with content revisions when we compared his drafts of this piece. From the first draft, which contained scattered paragraphs about his community service, he shifted to a more organized series of paragraphs. He began by framing the expectations for community service at his school, then explained his personal service to his grandmother (he cleaned her garage), and finally closed with how his involvement in community service could help him get into college. We were left with a dilemma for assessment of process: Neither his interview comments nor his written reflection suggest that he was able to explain his content revisions effectively, but there was some evidence of capacity to revise in his drafts. We are left uncertain of CFR's understandings of revision, his competence with revision, and his competence with the genre of analytic reflection. We are also left in the dark about when and where CFR might have gotten assistance on his work, whether from peers, his teacher, or his own self-reflection.

Generic analyses of revision without explication. Some students used the vocabulary of the writing process in ways that were largely external to any specific piece of writing. Some students might characterize revision or the role of peer response in ways that were consistent with their teacher's expectations but not articulate the links between their particular purposes for writing a piece and the processes they used to accomplish the writing. ASA, for example, explained the role of a "conference" in her classroom, and she cited generic questions that she asks "somebody" on these occasions:

If I get really stuck in school, I have a conference with somebody and ask them if you could help get this other part that I need to work on.... If they say, "yeah," then you go down and sit on the rug. I say, "I don't know what I should write next, can you help me on my next part?" (ASA, student writing, n.d.)

Examples from another student, BMK, reveal three different features of generic explanations. In the first example, BMK analyzes the processes she used when composing her persuasive letter, referring only to the assignment requirements, providing no convincing evidence that any of the steps or resources she used were helpful to her:

I started to do a page caled Planning my Peruasive letter. It helped me alot because I got to think about what I was going to write. Then I worte my rough draft and used the ideas from when I wrote on my Planning my Persuasive

Letter. Then I started to do my finlly copy and when I was done Mrs. Bentley said its great. Then I was happy.

In the second example, BMK uses the first person plural (*we*) to refer to typical revision practices in her classroom without reference to their application to specific revisions, even though she is asked to analyze her methods of revising a specific assignment: "What I did: *The first thing we did was to cluster our ideas for getting organized.*" Thus, her individual purposes and processes become submerged in a generalized pattern that everyone was to follow. This same example reveals a third pattern discernible only in situations in which portfolio readers have access to portfolios from the same classroom: We found an exact duplicate of BMK's "cluster" sentence in two other students' portfolios (BSL and BGL) for the same assignment; we believe that the description of process may have been written on the board and copied by the students.

When we talked with students about the roles of peers, we heard many generic reports of what *we* (or *you*) do when *we* respond to a peer's writing or revise on the basis of a peer's response. Notice how the following two examples (the first from BMJ, the second from DCN) reveal little about the ways either of these students was seeking and using peer response to accomplish particular purposes in his or her writing.

We correct the kids' work a lot. So when we do it, they just—sometimes we write notes what you could've done this better, and then they give us a grade and tell us why and all that.

It helps you [to] change something before you have to turn it in. Like you ask somebody, "Do you like this? And if you don't, what should I change?"

Reflections and responses that we classified as generic cannot be taken as evidence of lack of understanding; students who provided a generic response to our interview questions or to a teacher's reflective prompt might have elaborated further if either of these contexts had been better crafted to support their reflective analysis of the ways that they had revised a piece of text. This category serves to demonstrate that, without explication, generic responses may seem to be minimally analyzed repetitions of what teachers tell students about the value of revision or of peer response.

We were challenged by cases of generic explanations that did not explain a student's revisions, much as we were in our efforts to reconcile CFR's vague explanation with his more substantial content revisions (presented earlier). Consider the case of DDA's "five senses" assignment. In his first draft, DDA wrote an exceedingly brief description of an athlete named George "who played every sport." Although the assignment asked students to create images using the five senses, there was nothing in his initial single paragraph that would indicate his understanding of

this concept. In response to DDA, one student wrote: “You may want to use your senses. Maybe you want to tell about a certain event.” The second peer wrote: “You need to use your senses. Tell more about George playing a certain sport. What did George look like and what was his favorite sport.” DDA’s letter to Ms. Donner summarized these points, and his next draft included additional events and images—a description of George’s long jump event, a football game where George made the final touchdown, the smell of hot dogs in the air, the deafening noise of the crowd, and the touch of the athlete’s shirt soaking with sweat. In his final reflection on the piece, the student wrote:

My favorite warm-up is, “He was a great athlete.” The reason I liked this warm-up is because this was about a boy who loves to play sports and I think it kind of relates to me because I love to play sports. In this warm-up, I learned that responses really help your story a lot. The story also helped me put my five senses in and make the story interesting. My story was a para[graph] long and with the responses I made it a full page. (DDA, student writing, n.d.)

DDA mentioned that he made revisions that not only increased the length of his piece but helped him make newfound connections between George and DDA himself. Still, his reflection did not otherwise reveal what were substantive content changes reflecting his peers’ specific advice.

Based on our analyses of cases of generic student reflection, we came to recognize the potential of reflective writing to support a rater’s effort to produce a coherent evaluation. We do not argue that inclusion of a written explanation ensures a rater’s valid interpretation; a student’s analytic writing about process is a genre that itself requires practice and support and can itself be assessed (although this assessment is not a simple or straightforward process). Our point is that, although reflective writing cannot provide complete evidence of students’ understandings and competencies, such writing serves a critical function within the collection of portfolio artifacts (Figure 3).

Explications of revision processes. Students’ extended explanations of their revision processes provided compelling evidence of the ways they went about linking purpose with process as they attempted to revise and refine a piece of writing. In these explanations, students brought technical language to bear on analyses of specific changes that they had made in portions of a text:

[In our interview, CAA selected and read a piece that she thought needed further work.] I described too much, I think. And I really didn’t get on to the story. [We asked, “And what does that do to the reader, do you think?”] I think it really bores the reader, and while I’m writing, though, as a writer,

you don't really notice all these things ... then maybe when you look at it, you look back and see one page full of just description. Description, I need to cut something out.

[In an interview with another student, BSL, we asked, "Now is there a piece in your portfolio where you really feel like you worked really hard to change it?"] In the Belize letter—well I kept on wondering how to start this ... and I kept on changing my mind. In the end I came up with: "I think you should try hard to protect the rain forest. I did this because at the rate the rain forest is getting cut down, it will not exist in 8 years." 'Cause that is like really tense, because I think, "Oh my gosh this thing won't exist in 8 years" and when you're on, in the letter it says we can—"maybe we can not survive without it because it's considered the lungs of the world."

In these examples, students did not use the generic *we* and did not imply that certain steps in the process "had to" be done. Instead, they used *I*, making intensely personal claims for themselves as writers as they discussed specific revisions needed to keep a particular audience engaged. Thus, CAA realized that her lengthy descriptions could belabor her piece, and BSL realized that to persuade her formal and powerful audience (the King of Belize) she had to struggle with her writing, a process that made her feel "tense"—a word that displayed her engagement and effort. As BSL explained, she pointed to particular places where she accomplished her goals, remarking on the highly visual metaphor, *the lungs of the world*, as well as places where she labored ("Well I kept on wondering how to start this").

These explanations made explicit references to revisions in the drafts; the students provided talk about writing and about their understandings, demonstrating their attention to process and reflection. The close relation in these cases between students' interview reflections and the content of their revisions in the dated drafts shows the potential for young writers to close the loop of evidence shown in Figure 3. The artifacts of their writing are referenced in their reflections about the processes of their work; their reflections promise to influence the planning of future writing. Yet, it is critical to remember that these examples of clear explication on the part of the students came from our portfolio interviews—discussions that provided students with enhanced opportunities to reflect on their revision processes. We found no reflective writing bearing on these pieces in the students' portfolios and no audiotapes of reflective discussions with peers or teachers. Raters of these portfolios would not have had the benefit of our conversations with these children, and thus, without evidence of reflection, the raters had no access to students' reflective insights into their own writing processes. If portfolios are to be utilized as evidence of students' uses of writing processes, it may be critical that some portfolio entries include students' analytic reflection, whether written or recorded in some other way.

An aside: Peer response. We just reported that much of students' reflective writing about processes was either vague or generic in character. The same pattern was characteristic of peers' written responses; rarely were these responses explicitly relevant to the intent and techniques of a particular piece of writing. This finding concerned us in that in all four classrooms, peer response was the impetus for most revision. Because a writer's revisions emerge, in part, as resolutions of tensions between the writer's intent and a reader's suggestions, if the suggestions are not helpful, the writer may be left uncertain how to proceed.

We noted that patterns of response differed for weaker writers and for stronger writers. CFR, a writer who was struggling, received comments from peers that were elliptical at best: "To me I think you missed meeting [the expectations] just barely. Improve a few mistakes and you'll meet." When it came time to fill in a response to the question, "Based on the responses from my group critique, I will _____," CFR left it blank. In contrast, CAA, a very engaged writer, received responses suggesting feelings of intimidation. One peer responded as follows to CAA's letter to residents of Los Angeles: "Advise the writer what you think he or she might do to make it a better letter: *She should write clearly and not use high-class vocabulary words.*" In writing about how to improve the letter (a requirement), CAA resignedly said, "I can make it have more down to Earth vocabulary."

We came to see the blanks in process reflections of the weaker students and the grudging revisions of the stronger students as metaphors for the helplessness or frustration they may have felt when faced with responses that provided no suggestions for improvement of a specific piece. We became concerned that pedagogies that rely heavily on peer response may produce inequities in opportunities for helpful critique. Our focus here, however, is on assessability of portfolios that contain work guided primarily by peer response. How can raters ascertain a student's capacity to benefit from critique if the student rarely receives it?

Summary. From evidence in student interviews and portfolios, we constructed a framework for analyzing the ways that students reflect on their revisions of specific pieces of writing. We found that students rarely analyzed their writing processes orally or in writing in ways that provided compelling evidence of their understandings of revision. Although some conversations with students demonstrated substantive analysis of revision, for the most part students explained revision in generic terms as something they just had to do, rather than a process that would allow them to accomplish and enhance their purpose for writing a specific piece. We considered the relation between students' capacities to revise and students' understandings of the genre of explication both from the point of view of the student (what does the student know and what can he or she do?) and of the rater (what evidence does the rater have of either?), and we became increasingly aware of the

critical role of evidence provided by student's self-reflections on process. Without a student's own guiding analyses, we were left on our own to trace changes a student made across a series of drafts, a process that was complex and unlikely to be within the capacity of a rater pressed for time.

SUMMARY AND CONCLUSIONS

In this article, we reported findings regarding the creation and assessment of process evidence in CLAS language arts portfolios. Based on a small sample of teachers and students as well as on a portfolio assessment program that only trialed approaches to implementation, our findings were organized as frameworks and issues for productive discussion.

Two patterns emerged in our findings regarding students' opportunities to learn and to produce portfolio evidence. First, teachers' curricula varied in ways that provided students with quite different opportunities to learn about the writing process. Variations in the teachers' assignments reflected tensions between the romantic and the classical schools of composition (Hairston, 1986; see Figure 1). In the romantic view, the process of writing is flexible and responsive to the current piece, whereas in the classical view, the writing process is aligned with particular purposes and may be more procedural. Although the CLAS vision emphasized a balance of personal purpose with established forms, none of the teachers in our study fully represented that balance, and they tended to lean toward either end of the romantic-classical continuum. Thus, the curriculum in place in four classrooms varied in its alignment with the dimensions of learning.

Second, teachers varied in their approaches to documentation of students' writing processes in ways that provided students with quite different opportunities to produce and choose portfolio evidence. Portfolios more often contained evidence of editing (word changes, mechanics) than prewriting or revision. Artifacts revealing a student's efforts with the more complex processes that are the heart of the CLAS dimensions were often never created, or they emerged in forms that were difficult to analyze, such as cryptic notes or repeated drafts that required time-consuming content analyses. Some of the missing evidence emerged not from gaps in the curriculum but from teachers' pedagogic decisions to reduce the burden of process documentation, to minimize the risk that imposed requirements might stifle reflective processing. Missing evidence posed problems for our capacity to evaluate the portfolios. On the other hand, in classrooms in which process documentation was more consistent across assignments and portfolios, we questioned whether the documentation reflected students' writing processes or those of their teachers.

When students analyzed their processes in writing—the ways they drew on resources, developed early plans and drafts, solicited or interpreted input, or revised

pieces of writing—their reflections were often vague or generic and were not likely to reveal the ways they had composed or refined a piece of writing to accomplish a specific purpose. Most students either focused on editing or provided an explanation of revision that seemed to be a close repetition of their teacher's concerns for what the class "had to" do, without explication within a piece of text. We discussed how descriptive or analytic writing about process is a genre that itself requires practice and support. Indeed, students' superficial and limited responses to reflective prompts were evidence of both their understandings of writing processes and their understandings of reflective analysis.

What would constitute portfolio evidence that could support a rater's valid assessment of a student's uses of writing processes? Although the absence of crucial artifacts precluded assessment of process, the presence of artifacts was rarely a direct line to our understandings of a student's competence with or understandings of writing processes. We found it hard to imagine how a portfolio rater pressed for time could evaluate the complex of evidence that we found in many portfolios. In one classroom, for example, the teacher did not yet expect her young students to use the results of their brainstorming—would a portfolio scorer have time to reflect on the subtle differences between the existence of a particular step (web, sun) and the use of that step in composing? In another classroom, the teacher engaged her students in peer and self-assessment of many pieces of writing, expecting them to carry forward their learning from one assignment to the next—would a portfolio rater be able to track growth across pieces in the portfolios? How would a rater interpret how a student is growing in relation to the relatively generalized dimensions that made up the rubrics (e.g., "creativity")? Noticing patterns in the process evidence required our careful tracking and scrutiny, and time constraints on the rater may disallow substantive content analysis. We worried that the presence of brainstorming lists, organizing webs, written peer advice, multiple drafts, and written reflections could provide a rater pressed for time with an aura of effective processing in portfolios where closer examination might raise uncertainties about how these artifacts reflected the student's uses of them. Indeed, we sometimes found that the provision of multiple kinds of evidence—for example, artifacts from each phase versus written reflections—were inconsistent in ways that confounded rather than supported our evaluation. In this context, we grew increasingly appreciative of the critical evidence that could be provided by student's self-reflections on process: Who better to guide the rater through multiple drafts of pieces than the student himself or herself?

Our findings underscore the need to analyze closely how the presence or absence of particular kinds of evidence impacts particular rater judgments, how the need for evidence in the large-scale context may impact curriculum and pedagogy in the classroom, and how particular pedagogies in the classroom may support or impede the availability of evidence for large-scale portfolio assessment. We believe that

there can be far greater convergence and support of classroom and large-scale uses of portfolios than we found—predictably—in these four pioneering classrooms. We were not surprised that the portfolios were most likely to contain artifacts that support editing, on the one hand, and that, on the other hand, students were likely to focus their reflections on editing changes or on generic (“what we had to do”) characterizations of their procedures for content revision. The CLAS dimensions asked for reflective revision, yet the day-to-day written work in these field trial classrooms often centered on imposed processing requirements or surface editing. After all, the work of creating a coherent portfolio system was only just beginning.

How might teachers, students, and raters (who are teachers) build shared understandings of what is needed to ensure meaningful indicators of students’ writing processes? Teachers and their students need to know in advance how raters read and assess portfolios. Teachers (particularly those who may not also be raters) and students need opportunities to learn the assessment dimensions, perhaps through case examples of how raters have interpreted individual portfolios or through *think-alouds* of a rater’s reaction to individual pieces. At the same time, teachers may benefit from think-alouds and case examples that focus on the reasoning behind particular artifact use, specifics on how they model peer conferencing, or analyses of how specific assignments have the potential to achieve the CLAS balance between romantic and classical viewpoints. Understandings of cases and think-alouds can then be melded with individual philosophies and turned into curriculum and assessment practices that help students and teachers think about the evidence necessary for portfolio assessment of process and help researchers and raters interpret portfolio evidence appropriately.

We regret that the demise of CLAS made it impossible for us to follow classrooms over time. Further research is needed to understand how participants in a large-scale portfolio assessment program develop shared understandings of the ways that evidence of writing processes is considered in the scoring and how the programmatic needs for comparability of evidence can be reconciled with the personal needs of young writers, whose uses of processes will vary with the purposes and contexts of their writing.

ACKNOWLEDGMENTS

Both authors contributed equally to the work reported here. Maryl Gearhart is now at the University of California, Berkeley.

The work reported herein has received partial support from the Educational Research and Development Center Program, Cooperative Agreement R117G10027

and CFDA Catalog No. 84.117G, as administered by the Office of Educational Research and Improvement, U.S. Department of Education. The findings and opinions expressed in this report do not reflect the position or policies of the Office of Educational Research and Improvement or the U.S. Department of Education. We appreciate the detailed suggestions for revision of this article made by two anonymous reviewers and by Bob Calfee; the helpful guidance provided us throughout the larger project by ETS staff, including Joan Heller, Karen Sheingold, and Barbara Storms; and the wise counsel of the National Center for Research on Evaluation, Standards, and Student Testing's associate director, Joan Herman. We thank the teachers and students who generously shared their insights and their work with us.

REFERENCES

- Atwell, N. (1987) *In the middle: Writing, reading, and learning with adolescents*. Portsmouth, NH: Heinemann.
- Brown, M. W. (1949). *The important book* (L. Weisgard, Illustrator) New York Harper & Brothers
- Dyson, A. H., & Freedman, S. W. (1991). Writing. In J. Flood, J. M. Jensen, D. Lapp, & J. R. Squire (Eds.), *Handbook of research on teaching the English language arts* (pp 754-774). New York Macmillan.
- Erickson, F. (1986). Qualitative methods in research on teaching. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (pp. 119-161). New York: Macmillan.
- Gearhart, M., & Wolf, S. A. (1994). Engaging teachers in assessment of their students' writing: The role of subject matter knowledge. *Assessing Writing, 1*, 67-90
- Hairston, M. (1986). Different products, different processes. A theory about writing *College Composition and Communication, 37*, 442-452
- LeMahieu, P. G., Eresh, J. T., & Wallace, R. C. (1992) Using student portfolios for a public accounting *School Administrator, 49*(11), 8-15.
- LeMahieu, P. G., Gitomer, D. H., & Eresh, J. T. (1995) Portfolios in large-scale assessment Difficult but not impossible. *Educational Measurement: Issues and Practice, 14*(3), 11-16, 25-28.
- McGill-Franzen, A. (1993). *Shaping the preschool agenda: Early literacy, public policy, and professional beliefs*. Albany. State University of New York Press.
- Purves, A. C. (1992). Reflections on research and assessment in written composition *Research in the Teaching of English, 26*(1), 108-122.
- Sheingold, K., Heller, J., & Paulukonis, S. (1995) *Actively seeking evidence: Shujits in teachers' thinking and practice through assessment development* (Report No 94-04) Princeton, NJ: Educational Testing Service, Center for Performance Assessment
- Simmons, W., & Resnick, L. (1993). Assessment as the catalyst of school reform *Educational Leadership, 50*(5), 11-16.
- Spalding, E. (1995). The New Standards Project and English language arts portfolio. A report on process and progress. *The Clearing House, 68*, 219-224
- Thomas, W. H., Storms, B. A., Sheingold, K., Heller, J. I., Paulukonis, S. T., Nunez, A., & Wing, J. (1995). *California Learning Assessment System portfolio assessment research and development project. Final report*. Princeton, NJ: Educational Testing Service, Center for Performance Assessment

- Vermont Department of Education. (1990). *Vermont writing assessment. The trial year*. Montpelier, VT: Author.
- Vermont Department of Education. (1991). *"This is my best" · Vermont's writing assessment program, trial year 1990-1991*. Montpelier, VT: Author
- Wolf, S. A., Davinroy, K., & Gearhart, M. (1997, December). *"The clay that makes the pot": The loss of language in writing assessment* Paper presented at the meeting of the National Reading Conference, Scottsdale, AZ.
- Wolf, S. A., & Gearhart, M. (1997). New writing assessments: The challenge of changing teachers' beliefs about students as writers. *Theory into Practice*, 36, 220-230.
- Wolf, S. A., & Heath, S. B. (in press). Wondrous words: Young children's rewritings of prose and poetry. *The New Advocate*