C-R-E-S-S-T

National Center for Research on Evaluation, Standards, and Student Testing

Technical Report

You can view this document on your screen or print a copy.

- UCLA Center for the Study of Evaluation

in collaboration with:

- University of Colorado
- NORC, University of Chicago
- LRDC, University of Pittsburgh
- The RAND Corporation

"HOW DOES MY TEACHER KNOW WHAT I KNOW?"
Third Graders' Perceptions of Math, Reading, and Assessment

CSE Technical Report 395

Kathryn H. Davinroy, Carribeth L. Bliem,
and Vicky Mayfield
CRESST/University of Colorado at Boulder

February 1995

National Center for Research on Evaluation,
Standards, and Student Testing (CRESST)
Graduate School of Education & Information Studies
University of California, Los Angeles
Los Angeles, CA 90095-1522
(310) 206-1532

Copyright © 1995 The Regents of the University of California

The work reported herein was supported in part under the Educational Research and Development Center Program, cooperative agreement number R117G10027 and CFDA catalog number 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. Nor do these findings and opinions necessarily reflect the views of other project team members.

PREFACE

The current intense interest in alternative forms of assessment is based on a number of assumptions that are as yet untested. In particular, the claim that authentic assessments will improve instruction and student learning is supported only by negative evidence from research on the effects of traditional multiple-choice tests. Because it has been shown that student learning is reduced by teaching to tests of low-level skills, it is theorized that teaching to more curricularly defensible tests will improve student learning (Frederiksen & Collins, 1989; Resnick & Resnick, 1992). In our current research for the National Center for Research on Evaluation, Standards, and Student Testing (CRESST) we are examining the actual effects of introducing new forms of assessment at the classroom level.

Derived from theoretical arguments about the anticipated effects of authentic assessments and from the framework of past empirical studies that examined the effects of standardized tests (Shepard, 1991), our study examines a number of interrelated research questions:

1. What logistical constraints must be respected in developing alternative assessments for classroom purposes? What are the features of assessments that can feasibly be integrated with instruction?

2. What changes occur in teachers' knowledge and beliefs about assessment as a result of the project? What changes occur in classroom assessment practices? Are these changes different in writing, reading, and mathematics, or by type of school?

3. What changes occur in teachers' knowledge and beliefs about instruction as a result of the project? What changes occur in instructional practices? Are these changes different in writing, reading, and mathematics, or by type of school?

4. What is the effect of new assessments on student learning? What picture of student learning is suggested by improvements as measured by the new assessments? Are gains in student achievement corroborated by external measures?

5. What is the impact of new assessments on parents' understandings of the curriculum and their children's progress? Are new forms of assessment credible to parents and other "accountability audiences" such as school boards and accountability committees?

This report is one of three papers that were presented at the 1994 annual meeting of the American Educational Research Association and summarize current project findings.

Frederiksen, J. R., & Collins, A. (1989). A systems approach to educational testing. *Educational Researcher*, *18*(9), 27-32.

Resnick, L. B., & Resnick, D. P. (1992). Assessing the thinking curriculum: New tools for educational reform. In B. R. Gifford & M. C. O'Connor (Eds.), *Changing assessments: Alternative views of aptitude, achievement and instruction* (pp. 37-75). Boston: Kluwer Academic Publishers.

Shepard, L. A. (1991). Will national tests improve student learning? *Phi Delta Kappan*, *73*, 232-238.

"HOW DOES MY TEACHER KNOW WHAT I KNOW?" THIRD GRADERS' PERCEPTIONS OF MATH, READING, AND ASSESSMENT[1]

Kathryn H. Davinroy, Carribeth L. Bliem, and Vicky Mayfield
CRESST/University of Colorado at Boulder

Historically, efforts for school reform and research on such reform have tended to focus on institutions, state agencies of education, local districts, school boards, teachers, administrators, parents, and test scores describing student achievement. Reform efforts are fueled by public opinion showing outrage at the failures of public schools and demanding standardized tests to measure academic achievement of students (Elam, Rose, & Gallup, 1992). Interest in the relationship between assessment and instruction has led some camps for school change to look to new assessment procedures as a route to improving the learning of U.S. children (Resnick & Resnick, 1992). Yet scarce research is conducted *with* school children to find out the effects of these assessment and instructional reform efforts on children's understanding of what it means to go to school and learn. The voices of those most directly affected by school change are frequently missing from the literature. Thus, education reform becomes mainly the business of everyone but the student.

If reform is to make a difference, it must gain access to the perceptions of those it seeks to educate. Outcome measures focused on student learning, even classroom-based performance assessments, give us access to only a portion of the successes and challenges of the reform picture. If current professional definitions of reading and math emphasize meaning-making, then one object of educational reform is to assure that children are acquiring these understandings in school. Some research has shown a relationship between children's understandings of subject matter and their approaches to the tasks of reading (Borko & Eisenhart, 1986) and doing math. These tasks may be part of instruction as well as assessment. Speaking with students about their perceptions of subject matter and assessment gives us another window through which to view reforms in

---

[1] Paper presented at the annual meeting of the American Educational Research Association, New Orleans, April 1994.

assessment and instruction, as well as a view of changes in those student perspectives. Of what use is changing assessment and instruction to reflect new professional understandings of subject matter and learning if students continue to judge and to value their education in precisely those ways we seek to change? In addition to delving into student understandings, assessing our own progress in implementing reform may depend not only upon assessment measures that we created, but also upon the words and ideas of students.

Students have taken center stage in more recent, constructivist approaches to learning. Rather than acting as receptacles for knowledge, learners are actively involved in constructing knowledge from past experiences, from social contexts, through interactions with an environment that encourages the exchange and growth of knowledge (Bruner, 1977; Vygotsky, 1978). In this view, student understandings of what they are trying to learn are critical in guiding their efforts. This view that places the learner at the center suggests an education and assessment system that looks different from traditional schooling. One aspect of this shift from tradition is the need for alternative assessment practices—practices that address the process of learning, practices that involve the student in actively demonstrating her or his construction of knowledge. Performance assessments, then, are one strand of school reform that recognizes the centrality of the learner. Speaking with students about their school experiences with instruction and assessment can help us know better if new perspectives on learning have taken hold in the classroom.

As alternative forms of assessment make their way into schools and research, close attention should be paid to the effects of assessment reform on all parties involved, especially the learners. Are assessment reforms helping students internalize contemporary professional definitions of what it means to read and do math in school and other contexts? The present study, drawn from a larger project in which teachers developed and implemented performance assessments in their classrooms, investigates the children's perceptions of what reading and math are and how children understand their teachers' knowledge of them as readers and mathematicians. We ask: With the introduction of reading and mathematics performance assessments into the classroom, what are students' understandings of what it means to be a reader and mathematician and of what it means to demonstrate those abilities?

Conceptual Framework: Assessment, Subject Matter, and Children

One question to ask is whether *how* one is assessed influences one's ideas and definitions of *what* is being assessed. How students are asked to use and demonstrate their knowledge in a classroom may help mold how they perceive knowledge and knowing. These relationships among instruction, learning, and testing are part of the growing body of research surrounding assessment reform. Some studies illustrate relationships between assessment and instruction that show negative effects of standardized testing on instruction (Shavelson, Baxter, & Pine, 1992; Shepard, 1991; Smith, 1991). Other research shows that assessment influences, even drives, instruction—especially in high-stakes testing contexts (Lomax, West, Harmon, Viator, & Madaus, 1992; Shepard & Cutts-Dougherty, 1991). These studies raise questions about the sort of knowledge that is being emphasized in instruction as a consequence of formal assessment.

In light of these findings, some want to make assessment less narrow and more "authentic" (Wiggins, 1989)—that is, linked to more realistic learning goals—by using performances of competence to assess students' understanding of subject matter:

> Authenticity is more than face validity (Does the task look [for example,] like a reading task?) or curricular validity (Is the task consistent with the manner in which it is presented in the current curriculum?). A literacy assessment task is authentic to the degree that it resembles the way literacy is used in real life. It is not enough to be consistent with the curriculum, which itself may be disconnected from real-life literacy. A slightly less rigorous version of . . . authenticity . . . would be this: An assessment task is authentic to the degree that it promotes the use of literacy in the ways students are expected to use it in genuine communication acts. (Garcia & Pearson, 1991, p. 271)

If assessment can be shown to influence instruction, then, for example, literacy assessments that reflect real-world conceptions of reading and writing might guide instruction toward more meaningful representations and toward a reformed, "thinking curriculum" (Resnick & Resnick, 1992). The same holds true in mathematics. Some educational reformers tout a relationship between assessment and instruction as an effective way of shifting classroom practices from a skill-based philosophy to higher order thinking, as well as a way of providing educators with a more complex and complete picture of the learner. Implementing performance assessments might be one way to both initiate and measure shifts in instruction. Concomitant shifts in student perceptions of reading and

mathematics should also appear when instruction and assessment focus on thinking.

Subject area specialists in literacy and mathematics have supported the ideal of a thinking curriculum and of authentic assessment through various public statements of what constitutes reading and math and of what substantiates "excellence" in these areas. In *Becoming a Nation of Readers*, the Commission on Reading stated unequivocally: "The majority of scholars in the field now agree on the nature of reading: Reading is the process of constructing meaning from written text" (Anderson, Hiebert, Scott, & Wilkinson, 1985, p. 7). The Commission worried that when reading is taught and tested as decontextualized skills, students may very well learn and understand reading as discrete skills—not as an opportunity to make sense of text and the surrounding world. Reading as meaning-making has implications for instruction as well as assessment that include attention to strategies for fluency and extending comprehension into analytic and critical thinking. When reading is understood by teachers as meaning-making, both instructional methods and assessment change (Valencia, Hiebert, & Afflerbach, 1994). The effects of these changes on students' understanding of reading with the introduction of performance-based assessments are less well known.

In mathematics, the National Council of Teachers of Mathematics (NCTM) has formally adopted and published a set of curriculum and evaluation standards in order to ensure quality, indicate goals, and *promote change* in school mathematics (our emphasis, NCTM, 1989; NCTM, 1991). The standards document is built around a vision that includes learning to value mathematics, becoming confident in one's own ability, becoming a mathematical problem solver, learning to communicate mathematically, and learning to reason mathematically. The NCTM standards address not only what mathematics is and what it means to know and do mathematics, but also what teachers should do when they teach mathematics and what children should do when they learn mathematics. Significant changes in both instruction and assessment are required in order to implement the NCTM standards. Currently, researchers are studying what classrooms look like as teachers implement changes toward a meaning-centered approach in which students make conjectures, explain their reasoning, validate their assertions, and discuss and question their own thinking and the thinking of others (Ball, 1993; Lampert, 1990; Wood, Cobb, & Yackel, 1991). However, as

was the case with reading, rarely do we hear how these reforms are perceived by the students.

Research on student perceptions of learning and subject matter has been embedded largely in studies of metacognition (Baker & Brown, 1984) and attitudes toward *learning to* read or do math. Studies concentrating on student views have generally emphasized *attitudes* about subject matter and self-assessments of ability (Stipek, 1981; Stodolsky, 1985; Sturtevant, 1991). Others have focused on students' conceptions of reading as related to their experiences learning to read (Borko & Eisenhart, 1986). And, while some research has looked at students' perceptions of math as subject matter (Cobb, 1986; Desforges & Cockburn, 1987; Schoenfeld, 1989), few researchers have related these perceptions to assessment.

Student ideas about reading, math, and assessment should be more closely examined as standards for subject matter competence are disseminated across regions, districts, and schools. Calls for changes in definitions and standards related to mathematics and reading and parallel changes in expectations for learners have made educators aware of the need to reconceive their instruction and assessment practices. Outcome measures give us only one view of these changes. Speaking with students may give us a more complete view. Researchers, as scribes and interpreters of educational reform, must look to multiple sources of information, not just test scores and teacher reports of classroom practice, as markers of reform progress. We must hear more frequently from the students. What are their perceptions about the relationship between what they do in their classes and the methods their teachers use to determine their progress? These questions are valid even in classrooms where the change process is only beginning. This paper attempts to get at student understandings of these concepts by asking:

- What are students' perceptions of math and reading?
- What are students' perceptions of their teachers' assessment practices?

  What do students recognize as modes of assessment: reading aloud, author's chair, chapter tests, quizzes?

  What do students see as criteria for assessment: behavior, percent correct, effort, level of difficulty?
- What connections do students make among their definitions of reading and math, instruction, and assessment?

Methods: The Larger Project and a Context for Understanding the Children

We interviewed third-grade students from the classrooms of the three schools whose teachers took part in the larger project. The participating schools are located in working-class and lower-to-middle-class neighborhoods on the outskirts of a western metropolitan area. The district for this research was selected on three criteria: (a) an ethnically diverse student population, (b) a history of standardized accountability testing, and (c) district willingness to seek a two-year waiver from standardized tests in the three schools selected to participate. As principals and third-grade teachers volunteered to participate, it is possible that the communities in this project were predisposed toward school reforms efforts such as performance assessments (Shepard & Bliem, 1993). During the 1992-1993 school year that we worked with these teachers, they were exempt from the pressures of preparing for districtwide standardized tests. Rather, they spent the school year developing and implementing performance assessments.

The project began with each school-team of teachers choosing a reading and mathematics instructional goal for the semester that could be measured using performance assessments. With the aid of literacy and mathematics experts, the teachers agreed upon goals that they perceived as important for third graders to attain, and which were in keeping with district curriculum guidelines. The research team was then able to provide guidance and materials as the teachers implemented performance assessments aimed at measuring student progress toward those goals. In reading, the teachers chose to focus on fluency and comprehension with narrative text as their goals for the first semester. The literacy expert suggested and the teachers incorporated running records and summaries as assessment tools compatible with these goals. In the spring, as teachers included expository text in their reading instruction, both of these assessment tools were again used to gather data about fluency and understanding. Comprehension, as demonstrated with the written summary, was evaluated using a 4-point scoring rubric which the teachers negotiated during workshop sessions (see Appendix A). As school-teams, they established the criteria for each score and developed the language that would frame those criteria. They also outlined instructional activities that would introduce and illustrate the criteria for students. One teacher even reprinted the scoring rubric and attached it to every student-written summary as a way to remind students of the

standards for an excellent summary. Thus, summaries became a primary mode of instruction and assessment during the winter and spring months.

In math, the teachers elected to focus on the students' communication of their mathematical understanding. This goal took the form of explanations with the students being asked to explain the thinking that had led them to a solution of a math problem. This instructional goal necessitated the use of richer problem-solving tasks—tasks often provided by the mathematics specialist. As in reading, some teachers elaborated on explanations to a greater extent than others. One teacher wrote up a rubric similar to the summary rubric in reading as a way to capture both correct answers and good explanations. Others chose to grade responses separately along these two dimensions. In all classrooms, explanation of math problems was included in the mathematics instruction to some degree during our work with the teachers.

Because the entire project and treatment were structured according to subject matter, when we set out to talk with students, we devised an interview protocol that asked questions about mathematics and reading separately (see Appendix B). The interviews were counterbalanced so that one half began with reading questions and the other half with mathematics. For this study, we interviewed two students from each of the 13 classrooms at three times during the school year (fall, winter, and spring). Initial interviews were conducted in September 1992, with 26 students. We obtained a random sample of students, stratified according to teacher judgment of achievement, with 9 students categorized as high achieving, 8 students in the middle group, and 9 judged to be low achieving. The teacher judgments were revised for the winter sample, as the teachers thought they knew their students better by then. The second round took place in February 1993, with 20 new students and 6 repeats chosen randomly from the fall sample. The final round of interviewing occurred in May 1993, with, again, 20 new students and the same 6 repeats. We originally designed the study to interview the 6 repeating students each time in an attempt to assure that "student" was not a factor of any change in student perspective that we might see.

Based on preliminary analysis of the data, we did not observe much change, so the 6 repeats were analyzed as part of the whole sample for each round of interviews. This resulted in a total of 78 interviews. Of these 78, one complete interview and one reading portion were lost due to technical failure, and another

was omitted from the analysis due to a language barrier problem. This resulted in 75 interviews in reading and 76 in mathematics.

Our protocol included a number of probes enabling us to delve beyond single-word answers and to attempt to assure that students understood our questions. We conducted interviews during the regular school day in the hallways near students' classrooms. In a few instances, we interviewed students inside their classrooms while the rest of the class was busy away from the room. The interview protocol also included an invitation to students to retrieve items, papers, projects, and such—items that they felt their teachers used to determine how well they could read and do math. We sought to elicit students' perceptions of how their teachers understood their reading and math abilities and what evidence they believed their teachers consulted in order to form assessment judgments. However, given our knowledge of instructional and assessment changes that were going on in the classrooms, eventually we probed beyond what the students told us if they did not volunteer information related to the overall project. Thus, a subtle change in our protocol occurred between the fall and subsequent interview times: Upon recognizing the dearth of information from our fall interviews, we went to greater lengths to get the students talking in February and May by asking them specifically about summaries in reading and explanations in mathematics.

Analysis

Our analyses began with the authors transcribing all interviews. From these transcriptions, a careful reading of a sample drawn from the fall, winter, and spring interviews led us to a first-level coding system. The codes corresponded to our questions about understanding subject area, instruction, and assessment. We coded student responses as *definition* (of math or reading), as *instruction* (relating to their activities during the school day), and as *assessment* (their perception of their teachers' knowledge about them as learners). In describing our findings, student responses coded as "instruction" have been incorporated into discussions of definition and assessment because instruction tended to overlap heavily with both definition and assessment, and because these were our primary research categories. The three coding categories were further refined to differentiate between spontaneous responses and those requiring a greater degree of probing. We used this coding scheme systematically to categorize each student response to

interview questions and examined the coded responses for patterns related to our research questions.

Findings

Our findings are organized according to subject matter following the categories of definition and assessment. We present definition first because we expected that students' concepts of reading and math would be tied to their perceptions of assessment. In the discussion that follows, we examine this expectation.

Reading Findings and Discussion

The reading portion of the interview began by asking students if they saw themselves as readers. We envisioned that their answers to this and to follow-up questions about what they read and why they read would help us understand their definitions of reading. In response to the question "Are you a reader?" most replied "yes" but a few lower achieving students said "no" or "not really" (see Table 1). This pattern remained constant throughout the school year.

When asked why they read, nearly two-thirds of the students spoke of the enjoyment they receive from books. "Because it's fun to read . . ." "Because it's fun stuff. You can use your imagination and things" (see Table 2). These

Table 1

"Are you a reader?"

Response	Fall (n = 24)	Winter (n = 25)	Spring (n = 26)	Total (N = 75)
Yes	18	19	21	58
Sort of	1	4	2	7
Not really	3		1	4
No	1	1		2
Not asked	1	1	2	4

Table 2

"Why do you read?"
Responses of Students Who Call Themselves Readers (*N* = 68)

Category	*N*
ENJOY It's fun. I like it.	42
LEARN To learn new things and get information.	27
INTEREST It's interesting.	9
WORDS I can learn how to spell new words and what they mean.	10
GROW UP I will need the skill as an adult.	4
OTHER	6
NOT ASKED	1

Note. Students often gave more than one response, so total will reflect more than the number of students interviewed.

responses about how fun reading is or how much they like it were often combined with comments about how they read to learn. In fact, nearly one-half of the students suggested that reading helped them by teaching them new things.

> Well, sometimes I read to learn, sometimes I read to entertain myself, and I just like to read a good book.

> Because it is sometimes interesting and it teaches you things. And sometimes it makes me laugh.

In general, these students see the act of reading as a useful venture, either to gather information or to amuse themselves. Indeed, they seem to recognize and appreciate the meaning-making involved in reading.

Even those who refrained from labeling themselves as "readers" appeared to interpret reading as a meaningful and useful activity:

> Because I'm slow, like, when you read, if you don't get a word and I mean, I like to read, but I can't read very well.

. . . I have trouble with it. . . . I don't know any of the words.

And when asked what kept them from being readers, the students explained that their inability to read stemmed from word-level difficulties rather than from comprehension problems.

We sought to further understand the students' conceptions of reading by continuing with "I want you to think about what you do at school each day. When do you read?" We were hoping to hear about reading "period" as well as other instances of needing to understand text. Invariably students' initial response was to give us a time of day when reading instruction occurs in their classrooms: "At the end of the day" or "after Writer's Workshop." When probed, many students mentioned DEAR time, an acronym for "Drop-Everything-And-Read." They also spoke of reading at free time or whenever they completed their prescribed tasks. "We just, during reading time and sometimes when we have free time, like when we have finished our work and there's nothing to do, we read a book." But the end result was always the same: Regardless of when it happened, students read *books* and preferably chapter books. Only when explicitly probed did a few individuals admit to reading newspapers or magazines in the classroom.

Interestingly, a few lower rated students in the fall did talk spontaneously about having to read their math books. Story problems in math were, for this group, frequently mentioned when students were asked about reading at other times of the day. "Like instructions. Like to do a math paper, or something we don't know." In the spring, mid- and high-rated students also mentioned reading in other areas. "I have to read a math problem, I have to read what they ask you." In addition to subject-specific reading, a few children commented on other examples of reading during the school day. "We read our things that we put down, that we write. And we read stuff that she [the teacher] puts on the chalkboard."

We also asked students to reflect on any reading that they did outside of school, at night and on the weekends. As with school reading, the spontaneous reply was always books. When they read those books was usually at night or in the afternoons while playing school. When probed, some children would occasionally offer instances of reading the newspaper, magazines of interest, or comic books. Occasionally we heard about contexts other than book-reading:

And sometimes when we're going for a trip down the road I like to read the signs. To see where we're going and stuff.

> Sometimes I like to read like words that are just on other things like on a cup I like to read that or something. Or like on this machine [the tape recorder] or something I'd like to.

As was true in school, it was rare to hear any students mention reading as meaning-making other than with books, like with travel signs or other environmental print.

Once we had some understanding of what children perceived reading to be, we turned to assessment issues. Our second set of questions was aimed at eliciting students' understanding of their teacher's assessment in reading. We began by posing the question "Does your teacher know how well you can read?" We continued by asking how she knew, what they did to help her know, or what she thought if they weren't sure. Most thought that their teacher did know, and the number who said "no" or "I'm not sure" decreased during the course of the year (see Table 3). However, they were often unable to articulate any modes or criteria of assessment. More than one child commented that "teachers just know," suggesting in their tone that the answer was obvious. "She just sort of, you know, knows if I'm doing a good job." "She just thinks I do." They had real faith that their teachers knew everything they needed to assign grades in reading because, well, *she is the teacher*.

A few students told us that they didn't know if their teachers could tell how well they read because the teacher never told them. Several commented:

> Not really. . . . 'Cause I never hear her say it.

> Well, um, I think, I'm not sure. Um, she doesn't tell us how well we're reading or anything.

Table 3

"Does your teacher know how well you can read?"

Response	Fall (n = 24)	Winter (n = 25)	Spring (n = 26)	Total (N = 75)
Yes	15	21	24	60
No	2	1		3
Don't know	5	3	2	10
Not asked	2			2

It seems that these students perceived that they did not receive much feedback, or perhaps they did not recognize it if they did.

When we imagined this study, we thought that by offering students the chance to show us ways in which their teacher knows how well they are doing they would lead us through their classroom, pointing to their work on walls, their portfolios, and the author's chair, for example. Unfortunately, it seems these children rarely see or have time to digest evaluated work. More typically, the work that is returned to students is sent directly home as evidence to the parents of schoolwork. More than once a student waved off the idea that he or she might show us his or her work. "I took it home" or "I threw it away" were the common messages.

However, the vast majority of students were not distressed by a lack of evidence about what their teacher thinks. They assumed that their teacher knows what she is doing as well as what the students are doing without requiring support of the assumption. When we posed follow-up questions about *how* she knows, they often seemed perplexed, both that we would ask the question and perhaps because they didn't have an answer. Evidently, children seem unconcerned about their role in the assessment process though they are aware that they receive report cards on a regular basis.

Still, we persevered, encouraging students to think about how their teacher knows about their reading. We asked them to reflect on what they as students do that lets their teacher know. This often allowed the students to talk about instructional activities that constitute their reading time (see Table 4). For every instance that a child offered, we responded by asking if they thought their teacher used information from that activity to know how well they could read. We probed further by asking how the teacher knew they had performed well on the activity, or what she was looking for if she graded it.

The most often-cited reading instructional activity that also helped a teacher know how well a student read was that students read aloud to their teachers. Most of the children suggested this, directly or otherwise. They recognized this as a measure of fluency and familiarity with words. Even the child who didn't think his teacher knew how well he could read had as evidence that she didn't listen to him read. " 'Cause they don't listen to me read." A few students claimed that expression was the indicator being measured when they read aloud: "She's looking

Table 4

"How does your teacher know how well you can read?"

Response	Fall (n = 24)	Winter (n = 25)	Spring (n = 26)	Total (N = 75)
READ	15	18	17	50
RUNNING RECORD	0	5	7	12
QUESTIONS	14	14	7	35
WRITING	10	11	11	32
SUMMARIES	0	10	9	19
WATCH	6	4	3	13
GRADE	3	0	3	6
OTHER	5	9	11	25

READ	Student reads passages out loud for class or teacher.
RUNNING RECORD	Student reads a few pages from book while teacher listens.
QUESTIONS	Student answers questions about what she's read, either orally or on paper.
WRITING	Student composes written responses except summaries, including worksheets, answering questions, story maps, lit logs, diaries.
SUMMARIES	Student writes project summaries.
WATCH	Teacher watches students work to make sure they're behaving
GRADE	Teacher knows because "I'm in the highest reading group."
OTHER	Includes talking with teacher, quantity of material read.

Note. Students often gave more than one response, so column totals will reflect more than the number of students interviewed at each time.

for if I know when I should start a new sentence and when to slow down on my sentence . . . "

As the year (and the larger project) progressed, we heard the students talking about running records. Some, but not all, explained that this was how their teacher knew they could sound out big words. "She listens to us read and then she puts checks down for how good we can read and then the words that we have trouble with, she puts 'em down, too." By spring, most who had participated in a running record viewed it as a means of assessing their ability to figure out words. However, this was not because the teacher shared her notes from the record with them. One student said that even though her teacher didn't share the paper, she could see all the check marks on it:

> She, when we're reading this book about weather and she came around to my desk, and she was writing down all the words I didn't know. And most of them I did know. She put a check mark if I know the word. . . . Well, she doesn't show us, but you can see 'em while you're reading.

The degree to which students understood the criteria for the running record varied from one child to the next.

> She has like a blank piece of paper and she looks at you and she puts like little checks. I don't know why. Maybe if you get it right. She sometimes writes a word or something.

Few other assessment measures were mentioned spontaneously by students, especially at the beginning of the year. For this reason, we asked students if they did any writing about what they read. They talked about doing worksheets in reading time, usually skills-based lessons on contractions or spelling, or answering teacher-directed questions about what they had read. Later in the year, they included summaries and story maps in this list of written products. While they described the questions and summaries as tasks to tell "what the story was about," they often viewed the assessment of these same products as primarily based on handwriting and punctuation in addition to, or sometimes instead of, comprehension.

> They like check your handwriting and check and see if it's right and see if you spelled it correctly.

> If they really have a lot and the whole paper's filled and things. If we fill out the whole paper, if we have question marks and things and we have the exciting part of the story.

By the end of the year, many high-rated students (and some others as well) could recapitulate the scoring rubric developed by the teachers in project workshops as criteria for good summaries. Of course, the rubric was interpreted into the students' language. They talked about getting the questions right or wrong, based on what was really in the story. The words "understand" or "comprehend" were not part of their vocabulary when referring to the rubric:

> Well, if you would get a Thorough if you got like all good parts but not too much and then it's interesting or not. If you got a Solid, you would get, you don't have all the parts and it's a little interesting. And a Some you would get if you don't have very

> many parts and it's not interesting, and a Little is like you have no parts and you just didn't know.

Another child who outlined her understanding of the rubric proceeded to show the interviewer examples of her work. When the interviewer asked what she needed to do to make her summaries graded "3s" into "4s," the girl replied, "You need to have perfect spelling, periods go where they're supposed to."

A theme that ran throughout discussion of different modes of assessment (reading aloud, writing, teacher watching) was that behavior played a major role as a criterion. Students thought that *any* classroom reading activity could be judged upon how well they behaved while they were engaged. Obviously, classroom management is an issue for classes averaging 25 in number. Four years of schooling have led these students to conclude that their behavior counts no matter what the task. "She watches us read, and if we're not reading we got our name on the board. And you have to make up that time at recess."

Clearly, from their responses, these students generally recognize reading as a meaning-making task. They cite learning and enjoyment as reasons for reading. They even appreciate that tasks at school call on their abilities to comprehend what they have read. Yet, this message gets confused when they are formally assessed on their reading ability. Without specific criteria for meaning-making and practice with those criteria, students believe that assessment activities are often aimed at measuring their handwriting, at punctuation, at word recognition when reading out loud. The idea of text as meaning seems to get lost.

As discussed above, students rarely saw instances of their work with grades and comments on it. Until a child receives her paper with a score on it and explanation for that score, she may not internalize the rubric and apply it to her own work. During the course of the project, students did become familiar and sometimes comfortable with the scoring rubric for summary writing. It seemed the rubric was most internalized by higher rated students or by students who helped negotiate the rubric into their language.

Math Findings and Discussion

In turning to mathematics, we again tried to determine students' definitions of mathematics before asking questions about assessment. We probed for their definitions by asking them first about when, during the day, they used math, then whether they thought they were somebody who could do math, what were some of

the things they could do, why people needed to learn those things, and finally when they use math outside the classroom. Though we hoped for spontaneous responses that would include occasions in addition to "during math time" as a response to when they use math, they only admitted that math was used during other activities when we probed. When asked directly if they thought taking attendance or counting hot versus cold lunches involved math, several reluctantly agreed.

When we asked whether they thought they were somebody who could do math, almost everyone said they could. However, there were a couple of students (low to middle range) during each round of interviews who said "No," "I don't know," or "Sort of" (see Table 5).

We followed up this question with "What are some of the things you can do?" and had an overwhelming response pattern of a list of computational skills: "I can add, and take away and count"; and "like subtract. I don't know how to do times." In the spring this list grew longer with responses like "Well I can add, I can subtract, I can multiply, divide." Though the list of things to do in math expanded, it was still a list revolving around computation. In an effort to probe for a definition that went beyond computation, we questioned students about daily instructional time asking what they *did* during "math time." Some students revealed that during math time they occasionally played games, engaged in activities, and solved problems in addition to the worksheets and pages of math book problems that most students cited. However, they did not add these math games or problem-solving activities that they said they were doing in their classrooms to the lists of their abilities in math.

Table 5

"Are you someone who can do math?"

Response	Fall (n = 25)	Winter (n = 25)	Spring (n = 26)	Total (N = 76)
Yes	20	18	22	60
Sort of	1	3	1	5
No	1	1	1	3
Don't know	1		1	2
Not asked	2	3	1	6

Further into the interview we shifted the focus from their ability in math to the possible meanings of math for others. In general, these third-grade students felt people needed to learn computational skills "so we won't be stupid when we grow up and we won't learn nothing 'cause we won't have a job then," and because "it would help them be a better person." The occasional student was explicit about how people would use this type of math in a future activity. "So that they can balance their checkbook," or "When they get older they don't have to worry about counting on their fingers and if they give you a price you won't, like if you're at a bank, kinda like my mom is, you could add the money that the people need or want." Finally, they gave only examples of computation as uses of math outside of school if they gave anything other than homework or playing school. The students seem to have bought into the "an education will make you a better person" message, but, though they see math as useful, it is defined as computation.

The heart of our protocol tried to determine what were students' perceptions of assessment. We asked whether they thought their teachers knew how well they could do math, how the teacher knew, and what her criteria were. Most students were quick to respond that yes, their teacher knows how well they can do math. There were a few that said "I don't know" and only a couple "I don't think so" answers (one student qualified this response as being because it was still early in the year) (see Table 6).

The question of "how does she know" was not an easy one for the students to answer and sometimes got confused with how the student knew the teacher knew. For example, we heard responses such as "she says I'm one of the smartest students in the class" or "I'm in one of the high math groups." However, in getting further into the question, a few students mentioned things like "my teacher from

Table 6

"Does your teacher know how well you can do math?"

Response	Fall (n = 25)	Winter (n = 25)	Spring (n = 26)	Total (N = 76)
Yes	19	22	22	63
No	3			3
Don't know	3	3	4	10

last year told her" or "my mom told her," and most said something like " 'cause she grades my five minute tests" or "well, 'cause she looks at my papers." Most students seemed to have decided for themselves that their teachers use their tests and papers in math in order to determine how well they can do math (see Table 7). These tests and papers focused mainly on computation which seems to have helped develop their perceptions of mathematics as computation.

When we asked "What is she looking for on your papers?" or "How does she determine if you've done a good job?" the main criterion the students reported was right answers:

> Right answers, wrong answers, um, that's about it.
>
> If I get the answer right, she knows I was thinking hard.
>
> How much problems we get right and how much we don't.

Several combined the "right answer" criterion with handwriting, neatness, spelling, correct format, and behavior:

> That the answers are right and the handwriting's right and you borrow a lot.

Table 7

"How does your teacher know how well you can do math?"

Response	Fall (n = 25)	Winter (n = 25)	Spring (n = 26)	Total (N = 76)
PAPERS	13	18	19	50
TESTS	4	10	12	26
WATCHES	5	9	3	17
GRADES	2	1	0	3
OTHER	3	1	2	6

PAPERS	Math book problems and worksheets
TESTS	Timed tests as well as other math tests
WATCHES	The teacher watches the students as they work on their math, play games, do activities
GRADES	Report card grades
OTHER	Records from other schools/grades, mother told the teacher, student told the teacher, conferences

Note. Students often gave more than one response, so column totals will reflect more than the number of students interviewed at each time.

> She looks for if the answers are right, neat handwriting and if you have space in it. Like space between this way and down.
>
> . . . if all the words were spelled correctly and stuff.
>
> She's watching for somebody paying attention so they can do their work. And nobody being so stupid and funny.
>
> Oh. I think she thinks I've done very good because I'm quiet and I do my math and I get it done right on time.

Some students indicated that effort was yet another criterion for doing a good job: "Because she saw me working hard." In general, these students felt their teachers knew how good they were in math by assessing for "right" answers and by watching for logistics and behavior. Even when probed, no students mentioned their teacher checking for understanding.

Of these papers, the assessment tool that seemed to make the biggest impression on the students was the timed math-fact test. When we asked what does a math test look like, one student responded: "It has a hundred problems on it and you have to get as many problems as you can down in five minutes." The timed test also seemed to help create the impression that math is something you have to do quickly in order to be good at it. Several students made comments like: "And I'm doing really good on my timed test. I do them really fast. I'm getting really fast at them," or "Well, because I'm usually the first one done; I'm really good at it . . . I can get the number in my mind and just subtract or add it or multiply it real fast." The number of students who mentioned timed tests increased dramatically from 4 in the fall to 10 in the winter and 11 in the spring. Some teachers gave a lot of attention to timed tests according to the students: ". . . she looks at them and she says, well, we give people, we clap for the students that really improve by at least 10 or 15."

The larger project encouraged teachers to focus on creating assessments outside of the timed tests—a focus that shifted attention from computation to problem solving. As with summaries in reading, we worked with the teachers to create a performance assessment and rubric that asked students to write explanations of how they solved math problems. Because we had knowledge of the broader project and because we hoped students were moving beyond computation, we probed students to tell us if their teachers ever asked them to explain their answers. In this question we did see a shift between the fall and the winter/spring

interview times (see Table 8). In the later interviews, more students responded that they did explain their answers to math problems, at least occasionally.

There were some students who when probed could also talk about the scoring rubric for an explanation of a math problem.

I: OK. How do you know if you've given a good explanation?

S: Cause she'll usually give us a number like 4 or 3 or 2 or 1.

I: Do you know what those mean?

S: 1 means you did it a little bad, and 2 is kinda good, and 3 is good, and 4 is good.

I: Do you know what you have to do to get a 4?

S: You have to get the answer right and you have to explain it good.

These students also seemed to understand the contrast between the answer and the explanation of their solution. In reflecting on a 2 a student received on one paper, she said it meant "that I wasn't, that I didn't tell a lot about it, and I didn't have the right answer." However, sometimes what a student meant by explain was not always what we meant (nor what the larger project meant).

I: Does she ask you to explain your answers ever?

S: Yeah, sometimes.

I: What does she ask?

S: Like, uh, what's this answer, and I tell it to her and she says good and then I'm done.

Table 8

"Does your teacher ever ask you to explain your answer?"

Response	Fall (n = 25)	Winter (n = 25)	Spring (n = 26)	Total (N = 76)
Yes	8	17	12	37
Sometimes	6	5	7	18
No	8	2	3	13
Not asked	3	1	4	8

It did not appear that the students understood that the teacher might be using these explanations as indications of how well they were doing in math. Though explanations were being scored, the value of the explanations as a *mathematical* activity didn't seem to influence student understandings of how their teachers knew they could do math.

Our data suggest a picture of third-grade students who view math as arithmetic problems that have right and wrong answers. Students say their teachers assess them in ways that are consistent with this view by giving them tests and problems to do and by grading them mainly for correctness. Though the third-grade curriculum is heavy in arithmetic skills, reformers would hope that these skills could be combined with equal weight with problem-solving abilities. There is significant emphasis on getting third graders to "know their facts." However, the NCTM standards are trying to get away from this emphasis especially in the decontextualized timed-test setting.

Some Cross-Subject Similarities

Our findings about student perceptions regarding reading, mathematics and assessment support contentions that reform takes time if perceptions and understandings are going to change significantly (Anders & Richardson, 1992; Borko & Putnam, in press; Richardson, 1992). Interestingly, in addition to subject-specific findings, issues of communication and student access to graded work cut across subject areas.

Communication with third graders proved to be an occasional problem across subject matter. These problems seemed to be associated with our inability to ask a question that students understood in the way that we did. Communication difficulties were signaled by lengthy periods in transcripts in which student and interviewer engaged in a rapid back and forth exchange that didn't seem to arrive closer to a definitive response. When 20 to 40 lines of transcript are devoted to phrasing and rephrasing of the interview question, clearly a miscommunication is occurring.

One example of a communication problem exhibited in the reading interview occurred when we probed with the question "Does your teacher ever ask you questions about what you're reading?" We, as interviewers, thought we might hear instances of students retelling stories and telling us this demonstrated their comprehension. We imagined students would respond with questions like "What

happened to Bobby?" or "How was the problem solved?" But, while our third graders agreed that, yes, their teachers did ask them questions about their books, even with persistent probing students responded with examples most often in the form of "Do you like the book?" or "Is it the right level for you?" While these may be useful instructional questions, they did not tell us much about assessment, particularly assessment of understanding. Our parallel question in math, "Does your teacher ever ask you to explain your answer?" showed some similar though less consistent difficulties. When we probed students to talk about explaining problems, often they told us, "I tell it to her and I'm done." The differences between our expectations for responses and the actual student responses suggest that either we weren't communicating with the children, or teachers actually did not ask the sorts of questions we thought they were asking.

Another unexpected communication misunderstanding recurred when we probed students about how their teachers knew their reading abilities. Several answered by outlining how they knew their teachers knew they could read. The classic example was of the little girl who declared that her teacher knew she was an excellent reader "because I'm in the highest reading group." When probed about how the teacher knew to put her in this group, the student replied that it was because she was, in fact, a good reader. Students struggled with this question in math as well. When students tried to describe how their teachers knew how well they could do math, it was shifted or translated to how the *students knew* their teacher knew. Rather than describing a *process* by which teachers would form judgments, students were telling us how they knew their teachers' opinions, and this was based mainly on products or *results* of teacher assessment: I'm in the highest group; because I'm a good reader or math person.

Clearly at times, our students seemed to be trying to appease the interviewers. When the questioning went on long enough, they would list every instructional activity they could think of as possible vehicles for assessment. Literature logs, double entry diaries, the number of pages or books read were all offered as ways their teachers know how well they can read. Several explanations for the outcome are possible. The laundry list response could be an artifact of our interviewing style—students chose to list everything they do as a way to move beyond the question. Alternatively, the children could believe that such tasks—though very frequently ungraded—are used by their teachers to know how well they can read. In fact, talking with the teachers about their assessment

practices, several mentioned the "gut feeling" they have about their students just from observing them every day (Borko, Flory, & Cumbo, 1993).

The research literature on student thinking and perceptions prepared us somewhat for the eventuality of communication problems with third graders. Indeed, the difficulties of interviewing young children may contribute to the paucity of studies about their perceptions and understandings of subject matter as related to instruction and assessment. Research on young students' thinking and perceptions may well be hampered by the fact that children have not yet developed a causal understanding of the relationship between ability and achievement. According to Wittrock (1986), "children's concepts of the causes of their successes and failures develop from a relatively undifferentiated state to a more analytic conception of the relations among ability, effort, and achievement . . . at about 7 to 8 years of age they distinguish these concepts from one another, and causally relate effort, but not ability, to achievement" (p. 304). Asking students if their teachers know how well they can read and do math assumes that students understand that teachers are concerned with achievement and go through an explicit process for judging student performance. It is possible that scholastic achievement may be a nascent concept for third graders. But this possibility should not deter others from pursuing students' ideas of school subjects. Only through open-ended conversations can we more deeply probe and understand whether students see reading as meaning-making and mathematics as broader than computation.

Another cross-subject phenomenon involved our repeated efforts to get students to share work they felt their teachers used to assess them; we were often met by blank looks or responses that indicated they weren't sure what we meant. Very rarely did our question about graded work that their teacher would use to determine how well they could read or do math result in physical artifacts—like a scored summary or math explanation. The majority of students did not show us any work. One inference that might be drawn from this cross-subject similarity is that students do not seem to get to see enough of their work after it is graded to understand the assessment value that it has for their teachers, nor to understand how classroom performance gets translated into teacher assessment. Papers get thrown away, filed away, or sent home in folders to parents. Typically the path of student work moved from student to teacher to parents or to a file, and rarely was work returned to students for individual and/or class discussion. One teacher told

us that in an ideal world, papers would be graded and returned to students quickly, and would be followed up with discussion because students need the feedback in a timely way. Time-in-the-day continues to act as a major constraint in trying to vary the path of assessment artifacts and results.

Some Closing Thoughts and Implications

Our recommendations from this study begin with where we started: students' conceptions of reading and mathematics. Clearly, from their responses, these students recognize reading as a meaning-making task. They cite learning and enjoyment as reasons for reading. They even appreciate that instructional tasks at school call on their abilities to comprehend what they've read. Yet, this recognition gets distorted when they are assessed on their reading ability. These students believe that assessment activities are often aimed at measuring their handwriting, punctuation, and expression when reading out loud. The idea of reading as meaning shifts to reading as word identification. Responses to mathematics questions took a slightly different tack. Whereas students described reading as meaning-making and reading assessment as skills-based, in mathematics they demonstrated consistency across definition of math and assessment of math. In both categories, math is arithmetic problems that have right and wrong answers. Teachers know how well students can do math as a function of how many right and wrong answers they produce—not according to their thoughtfulness in problem solving.

With respect to project-specific reading and math tasks, students were aware of these tasks though few saw them as ways their teachers were assessing them. In reading, teachers purportedly used summary writing to assess comprehension, but students maintained a skill-based stance by explaining that their teachers looked for spelling and other mechanics in assessing summaries rather than providing the "gist" of a passage as called for by literacy experts (Palincsar & David, 1991; Pearson & Fielding, 1991). Our student perceptions of summary writing align with findings of some of the teachers' understandings of summary as reported in an early project study. Findings from the fall semester in one school, suggest that teachers agreed to use summaries to assess comprehension but found themselves focusing on using summary as an end in itself (Borko, Davinroy, Flory, & Hiebert, 1994).

In math, the larger project introduced into classrooms a number of math activities with the intent that they would be used for both instructional and assessment purposes. While students occasionally described "explaining" their answers as one way their teachers knew how well they could solve a problem, for the most part students saw these math activities as opportunities to produce neat work and to work hard. Student definition and understanding of assessment continue to follow typical ways of perceiving math—as computation. Getting away from math being simply computation is one of goals of the NCTM standards, and some say that students will learn what is valued by what is assessed. If the students don't recognize that their teachers assess them on their understanding of mathematics, how can they come to value understanding? If we want students to better understand and value the meaning-making and communication involved in mathematics, we need to find ways to make these connections more explicit for the students.

With respect to reading, it seems teachers need to continue to work toward authentic performance assessments that extend and build on students' existing understandings of reading. In math, the task is more challenging. That students continue to perceive math as arithmetic problems and that teachers continue to support this perception with assessments that align with that view reflect the problems in implementing the fundamental changes called for by NCTM—to shift the focus from skills to mathematical thinking.

As discussed above, in both reading and math students need to see instances of their work with grades and comments on it. Too often, it seems, the papers are handed in, never to be seen again unless it's in the folder for parents. And it doesn't appear to be enough to work together as a class to derive scoring criteria. Until a child receives her paper with a score on it and an explanation for and discussion of that score, she may not internalize the rubric and apply it to her own work. Our findings raise some interesting questions about how students relate their concepts of subject matter to assessment practices: Should they be made aware of every instance that will be scored? Would this awareness, like public scoring criteria, help them, especially low-achieving students, perform more to their potential? How can we help to initiate a culture in the classroom where students realize that everything they do helps their teacher know how well they can read and do math, so that their teacher can aid them in becoming better readers and mathematicians?

Reform efforts arise from every aspect of educational research. When assessment reform can help spur reform in instruction, the many facets of education may beable to work together for general and lasting reform. But it is also the learners who must speak to these reforms. As efforts to demonstrate new and better instruction accelerate, students must be consulted more frequently as a gauge for reform success.

This paper is one beginning in trying to hear from the students about educational reform efforts, about how they understand their experiences in school, about how they form understandings of what it means to be a reader and mathematician. If it can be shown that performance assessments as an alternative to standardized and limited-format testing are effective in measuring student achievement; if it can be shown that instruction developed from performance assessments involves the whole classroom community in learning and assessing; and if it can be shown that students exposed to these new ways of assessing and instructing are constructing new meanings of subject matter—meanings that break with traditional and limited modes of understanding what it means to be literate and do math; then perhaps calls for a thinking curriculum may be closer to being realized.

References

Anders, P., & Richardson, V. (1992). Teacher as game-show host, bookkeeper, or judge? Challenges, contradictions, and consequences of accountability. *Teachers College Record*, *94*(2), 382-396.

Anderson, R. C., Hiebert, E. H., Scott, J. A., & Wilkinson, I. A. G. (1985). *Becoming a nation of readers: The report of the Commission on Reading*. Washington, DC: National Institute of Education, U.S. Department of Education.

Baker, L., & Brown, A. (1984). Metacognitive skills and reading. In P. D. Pearson (Ed.), *Handbook of reading research* (pp. 353-394). New York: Longman.

Ball, D. L. (1993). Halves, pieces, and twoths: Constructing and using representational contexts in teaching fractions. In T. P. Carpenter, E. Fennema, & T. Romberg (Eds.), *Rational numbers: An integration of research* (pp. 157-196). Hillsdale, NJ: Lawrence Erlbaum Associates.

Borko, H., Davinroy, K. H., Flory, M., & Hiebert, E. H. (1994). Teachers' knowledge and beliefs about summary as a component of reading. In R. Garner & P. Alexander (Eds.), *Beliefs about texts and instruction with text* (pp. 155-182). Hillsdale, NJ: Lawrence Erlbaum Associates.

Borko, H., & Eisenhart, M. (1986). Students' conceptions of reading and their reading experiences in school. *The Elementary School Journal*, *86*(5), 589-611.

Borko, H., Flory, M., & Cumbo, K. (1993, April). *Teachers' ideas and practices about assessment and instruction*. Paper presented at the annual meeting of the American Educational Research Association, Atlanta.

Borko, H., & Putnam R. T. (in press). Learning to teach. In R. C. Calfee & D. C. Berliner (Eds.), *Handbook of educational psychology*.

Bruner, J. (1977). *The process of education*. Cambridge, MA: Harvard University Press.

Cobb, P. (1986). Contexts, goals, beliefs, and learning mathematics. *For the Learning of Mathematics*, *6*(2), 2-9.

Desforges, C., & Cockburn, A. (1987). *Understanding the mathematics teacher: A study of practice in first schools.* London: The Falmer Press.

Elam, S. M., Rose, L. C., & Gallup, A. M. (1992). The 24th annual Gallup-Phi-Delta-Kappan Poll of the public's attitudes toward the public schools. *Phi Delta Kappan*, *74*(1), 41-53.

Frederiksen, J. R., & Collins, A. (1989). A systems approach to educational testing. *Educational Researcher*, *18*(9), 27-32.

Garcia, G. E., & Pearson, P. D. (1991). The role of assessment in a diverse society. In E. H. Hiebert (Ed.), *Literacy for a diverse society* (pp. 253-278). New York: Teachers College Press.

Hiebert, E. H. (Ed.). (1991). *Literacy for a diverse society: Perspectives, practices, and policies*. New York: Teachers College Press.

Lampert, M. (1990). When the problem is not the question and the solution is not the answer. Mathematical knowing and teaching. *American Educational Research Journal*, *27*(1), 29-63.

Lomax, R. G., West, M. H., Harmon, M. C., Viator, K. A., & Madaus, G. F. (1992). *The impact of mandated standardized testing on minority students*. Boston, MA: Boston College, Center for the Study of Testing, Evaluation, and Educational Policy.

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: Author.

National Council of Teachers of Mathematics. (1991). *Professional standards for teaching mathematics*. Reston, VA: Author.

Palincsar, A. S., & David, Y. (1991). Promoting literacy through classroom dialogue. In E. H. Hiebert (Ed.), *Literacy for a diverse society. Perspectives, practices, and policies* (pp. 122-140). New York: Teachers College Press.

Pearson, P. D., & Fielding, L. (1991). Comprehension instruction. In R. Barr, M. L. Kamil, P. B. Mosenthal, & P. D. Pearson (Eds.), *Handbook of reading research* (Vol. 2, pp. 609-640). New York: Longman.

Resnick, L. B., & Resnick, D. P. (1992). Assessing the thinking curriculum: New tools for educational reform. In B. R. Gifford & M. C. O'Connor (Eds.), *Changing assessments: Alternative views of aptitude, achievement, and instruction* (pp. 37-75). Boston: Kluwer.

Richardson, V. (1992). The agenda-setting dilemma in a constructivist staff development process. *Teacher and Teacher Education*, *8*, 287-300.

Schoenfeld, A. H. (1989). Explorations of students' mathematical beliefs and behavior. *Journal for Research in Mathematics Education*, *20*, 238-255.

Shavelson, R. J., Baxter, G. P., & Pine, J. (1992). Performance assessments. Political rhetoric and measurement reality. *Educational Researcher*, *21*(4), 22-27.

Shepard, L. A. (1991). Will national tests improve student learning? *Phi Delta Kappan*, *72*, 232-238.

Shepard, L. A., & Bliem, C. L. (1993, April). *Parent opinions about standardized tests, teachers' information and performance assessments*. Paper presented at

the annual meeting of the American Educational Research Association, Atlanta.

Shepard, L. A., & Cutts-Dougherty, K. (1991, April). *Effects of high stakes testing on instruction*. Paper presented at the annual meeting of the American Educational Research Association, Chicago.

Smith, M. L. (1991). Put to the test: The effects of external testing on teachers. *Educational Researcher*, *20*(5), 8-11.

Stipek, D. J. (1981). Children's perceptions of their own and their classmates' ability. *Journal of Educational Psychology*, *73*, 404-410.

Stodolsky, S. (1985). Telling math: Origins of math aversion and anxiety. *Educational Psychologist*, *20*(3), 125-133.

Sturtevant, E. G. (1991). Reading perceptions of urban second graders. In T. V. Rasinski (Ed.), *Reading is knowledge* (pp. 63-69). Pittsburgh, PA: College Reading Association.

Valencia, S. W., Hiebert, E. H., & Afflerbach, P. (Eds.). (1994). *Authentic reading assessment: Practices and possibilities*. Newark, DE: International Reading Association.

Vygotsky, L. S. (1978). Interaction between learning and development. In L. S. Vygotsky, *Mind in society* (pp. 79-91). Cambridge, MA: Harvard University Press.

Vygotsky, L. S. (1992). *Thought and language* (A. Kozulin, Trans.). Cambridge, MA: MIT Press.

Wiggins, G. (1989). Teaching to the (authentic) test. *Educational Leadership*, *46*, 41-46.

Wittrock, M. C. (1986). Students' thought pieces. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 297-314). New York: Macmillan.

Wood, T., Cobb, P., & Yackel, E. (1991). Change in teaching mathematics. A case study. *American Educational Research Journal*, *28*, 587-616.

APPENDIX A

Scoring Rubrics for Narrative and Expository Summaries

For Narrative text:	For Expository text:
4 (Thorough) — Includes the setting, the characters, and plot. Told in an interesting way.	4 (Thorough) — Organized. Includes main idea and some support. Written in student's own words.
3 (Solid) — May be missing some important parts of the story. May have too many details.	3 (Solid) — Not completely organized, but still flows. Includes main idea and some support. May have copied some phrases directly from the text.
2 (Some) — Doesn't have the most important parts. May have some wrong information. Not told in an interesting way.	2 (Some) — May not have main idea, but includes some details OR may have main idea but includes wrong information. May include wrong supporting details.
1 (Little) — Doesn't make sense. Not enough information is given.	1 (Little) — Includes a lot of incorrect information. Focus is on unrelated details or only on supporting ideas.

APPENDIX B

Student Interview Protocol

Reading/Math

1. First I'd like to ask you some questions about reading. Are you a reader?

If yes, what do you read?
Why do you read? If "because I like it", why do you like it?

If no, what keeps you from being a reader?

2. I want you to think about what you do at school each day. When do you read?

Probe (if child focuses on formal reading time): Do you read things at other times?
If yes, what do you read?
If no, probe by asking about other subject areas like "what about social studies?", and then offering examples such as science books, magazines, newspapers.

3. Does Ms. _____ know how well you can read?

If no, what does your teacher think? Why do you think she doesn't know how well you can read (or why is that)?

If yes, how does your teacher know?

Can you give me an example? You can go to your desk to get something, or take me there to see things. What kinds of things is she looking for?

For example, "Can you describe (that activity) for me?", "What does your teacher do/look at?", "Can you show me (this activity)?"

For conferences, What do you do? What does your teacher do? What does she write? Do you see/hear what she's looking for? Does your teacher ask questions? About what?

How does your teacher decide if you've done a good job (for each mode of assessment)?

Do you help your teacher decide if you've done a good job?

Probe: Look around the classroom. Is there anything else?

4. Now think about what you do after school, at night, and on the weekends. Can you think of times when you read outside of school?

Probe for "what do you read (for each time mentioned)?"

If examples focus on school/homework, probe by offering examples: reading to siblings, read TV Guide, etc.

5. Now let's go back to school and talk about math. Think about your school day and tell me when you do math.

What do you do during that time?

Probe (if student focuses on math time only): Are there any other times during the day when you use math?
If yes, what do you do?
If no, probe by offering examples of lunch count, attendance, etc.

6. Do you think you are someone who can do math?

If yes, what are some of the things that you can do?
Given examples, ask "Why do people need to learn these things?"

If no, what keeps you from doing math?

7. Does Ms. ____ know how well you can do math?

If no, what does your teacher think? Why do you think she doesn't know how well you do math (or why is that)?

If yes, how does your teacher know?

Can you give me an example? You can go to your desk to get something, or take me there to see things. What kinds of things is she looking for?

For example, "Can you describe (that activity) for me?", "What does your teacher do/look at?", "Can you show me (this activity)?"

For papers/worksheets, what do the questions look like? What does the teacher do? Do you know what she's looking for? What? Does she ask you questions? About what?

How does your teacher decide if you've done a good job (for each activity)?

Probe: Look around the classroom. Is there anything else?

8. Now think about after school, at night, and on the weekends. Can you describe for me a time when you use math outside of school?

Probe for "What do you do (for each time mentioned)?"

If examples focus on school/homework, probe by offering examples: allowance, pages read in book, buying things.

9. Is there anything else you want to tell me about reading or math?