CROSSING THE BORDERS AGAIN: CHALLENGES IN COMPARING QUALITY INSTRUCTION IN MATHEMATICS AND READING

Anna O. Graeber
Kristie K. Jones
Marilyn J. Chambliss

University of Maryland

Paper Presented at the Presidential Invited Symposium, Looking in Classrooms: Again, Annual Meeting of the American Educational Research Association San Francisco, CA, April 8, 2006. Direct all correspondence to A. Graeber or M. Chambliss at 2311 Benjamin Building, College Park, MD 20742. The work reported herein was supported by the Interdisciplinary Educational Research Initiative (IERI # 0115389), a combined effort of the National Science Foundation, the U.S. Department of Education, and the National Institutes of Health. The opinions expressed in this manuscript are our own and do not reflect the positions and policy of the National Science Foundation, the U.S. Department of Education, or the National Institutes of Health

Abstract

Although many studies have looked at the teaching of mathematics or at the teaching of reading, few have looked at teaching in both subjects and attempted to discuss characteristics of quality instruction across the two domains. The authors reflect on their work on such a study, drawing comparisons with the difficulties they faced and those faced by teachers attempting to construct interdisciplinary courses. Using the construct of cognitive demand, they explore the tension between pursuing general pedagogical principles versus the unique aspects of teaching that characterize effective teaching in one of the subject areas. They discuss the struggles faced in data collection, terminology used, and the comparability of research bases. Their project's findings suggest that the level of cognitive demand exhibited in the tasks teachers pose and the responses and work of students are similar in mathematics and reading. However the cognitive demand of content was found to be higher in reading than in mathematics. Such findings raise issues related to teacher education as well as research. Should teacher education incorporate information about quality instruction across disciplines, and, if so, how? Or, if quality instruction across such subjects differs, what are implications for the current expectations for elementary teachers to teach across subject matter disciplines?

Crossing The Borders Again: Challenges In Comparing Quality Instruction In Mathematics And Reading

The purpose of this paper is to explore challenges faced in seeking to characterize and compare high quality teaching for reading and mathematics, arguably the two most important "subject matters" in elementary school. In many elementary schools, the same teachers are responsible for both reading and mathematics instruction, and, of course, the same children learn them. And yet, reading and mathematics may be so different both to teach and to learn that the differences are more salient and meaningful than the similarities. Different professional groups have developed standards for the teaching and learning of reading and mathematics, and for the most part, different researchers have explored the relationships between high quality teaching and student outcomes in the two subject matters. As Stevens, Wineburg, Herrenkohl and Bell (2005) have argued, the assumption of knowledge domain specificity has led to a paucity of recent work that crosses subject matter boundaries. In this paper we briefly discuss studies of quality teaching that have included mathematics and reading, the challenges faced in studying teaching across subjects, and some ways in which high quality teaching in the two subjects appears to be similar. In these discussions we draw on the literature as well as our own work.

Previous Comparisons of the Teaching of Mathematics and Reading/Language Arts

As reported above, there is little work that considers teaching across any two subject areas. Stodolsky's (1988) study of elementary mathematics and social studies is probably the most well-known recent exception. Likely the most frequently recalled work that included both mathematics and reading were the 1960s era process/product studies. These relatively large-scale studies involving classroom observations were often funded by the National Institutes of Education and frequently used achievement test data facilitated by mandates to use such measures in federally funded programs such as "Follow Through" and "Head Start." While both mathematics and reading programs at both the primary and intermediate grades were studied, the results were often mixed and the studies have been widely criticized. The concerns voiced were about methodology as well as reliance on standardized tests as the measure of achievement (Erickson, 1986, p. 133).

Reports of this work and reports of coding schemes sometimes overlooked subject matter differences. For example, Good and Brophy's 1984 edition of *Looking in Classrooms* made little mention of subject matter differences other than differences in allocated time across subjects and teachers' penchant for allocating more time to subjects they liked (p. 31-33). Rosenshine's (1979) brief summary made few distinctions between mathematics and reading instruction and the summaries of many of the cited studies were devoid of any mention of the subject matter.

Given the limitations of the process/product research and the "methods experiment" studies that followed (Medley, 1979, p.14), conclusions of this work must be made with some caution. But there was evidence from the studies indicating that teaching varied across subjects. Evertson, Anderson, Anderson, and Brophy's (1980) study of junior high mathematics and English classes found that the teachers of these subjects used different instructional patterns, including different choices for use of class time, different usage of textbooks, and different amounts and types of questioning. Evertson et al. (1980) also reported differences in the correlations that a given instructional activity had on student achievement in the two subject areas. While in mathematics classes there was a positive relationship between the proportion of questions that required explanations and student achievement, this was not the case for English classes. Interestingly, they also found differences in how students rated mathematics and English teachers. Academically demanding mathematics teachers tended to be rated positively by students whereas academically demanding English teachers tended to be rated negatively. This may suggest that students too have different views of good teaching for different subject areas.

In elementary school, the same people typically teach mathematics and reading/language arts, the analog to English in middle school. Two recent studies focusing on elementary school teachers in high poverty schools arrived at a similar conclusion despite employing different methodologies (Knapp, Shields, & Turnbull, 1995; Rowan, Correnti, & Miller, 2002). Even when the two subject matters are taught by the same people, the type and effectiveness of instruction can differ. Knapp and colleagues (1995) conducted observations of grade 1-6 classrooms (some over two years), identified patterns, assigned codes to teachers accordingly, and drew conclusions about the outcomes. For Knapp et al. (1993), outcome measures included specially developed tests of writing and of problem solving, appropriate levels of the CTBS in mathematics, and parts of the Woodcock for reading. Rowan and colleagues (2002) reanalyzed data from *Prospects: The Congressionally Mandated Study of Educational Opportunity*. The *Prospects* study provided student data from standardized tests in reading and mathematics for two cohorts over three years (grades 1-3 and grades 3-5) and yearly surveys of teachers' practices and content coverage. Rowan et al. (2002) conducted statistical analyses to identify relationships among instruction and student achievement gains. Despite these methodological differences, both analyses led to the same conclusion across reading and mathematics. According to Knapp et al., "…what teachers in our sample did in one subject area reveals relatively little about what they did in another" (1992, p. 9). Rowan et al. explained that "a given teacher varies in effectiveness when teaching different academic subjects" (2002, p. 1533).

Both studies add to our understanding of the nature of the differences. Knapp and associates (1995) distinguished among teachers according to whether they engaged in instruction that departed substantially from conventional practice or provided "alternative" practices. Specific alternative practices differed for the two subject areas, and the authors did not directly contrast subject areas in this analysis. They did however consider the alternative practices in the two subject areas at a more abstract level. Alternative practices were considered as those which gave emphasis to meaning and understanding, embedded skills in context, and facilitated connections among subjects and between subjects and life outside of school for both reading and mathematics instruction. The analysis revealed that only a small number of teachers were engaged in instruction that departed substantially from conventional practice in more than one subject area (Knapp et al., 1992).

The *Prospects* reanalysis did highlight differences between mathematics and reading using two types of analyses (Rowan et al., 2002). First, the reanalysis contrasted the effect of whole class instruction with individualized instruction on student academic growth in mathematics and reading. Noting that process-product research suggested a positive relationship between amount of active teaching and student achievement [see Brophy and Good’s (1986) review of process-product research], they chose whole class instruction as a proxy for active instruction in contrast to individualized instruction, which would not involve active teacher involvement. The reanalysis revealed that the amount of time that teachers reported employing whole group settings was related to student academic growth in both reading and mathematics. Interestingly, the amount of time that students spent in individualized settings was negatively related to student achievement in reading, as would be expected from previous research, but not in mathematics. Rowan and associates (2002) did not speculate on reasons for this difference between the two subject areas.

A second analysis focused on the relationship between cognitive demand and student achievement in the two subject areas. Rowan and colleagues (2002) constructed scales for mathematics and reading based on the cognitive demand of the content covered, as reported by teachers. For students who progressed from grades 1-3 during the study, cognitive demand was related to academic growth in reading, but not mathematics. The authors did not speculate on reasons for these differences. . Cognitive demand was related to academic growth in mathematics for the 3-5 cohort, but lack of items had made it difficult to construct a scale for the 3-5 cohort in reading.

Both of these more recent studies address criticisms levied at the earlier process-product work. Both of them took subject matter context into account. Knapp and associates (Knapp & Marder, 1992) developed different observation codes and ordinal scales by which to sort and order teachers for mathematics and reading. Rowan and colleagues (2002) developed different ordinal scales by which to order teacher responses to content questions for both subject areas. Furthermore, both of them privileged instruction that came from a cognitive constructionist paradigm (Knapp et al., 1995; Rowan et al., 2002). Knapp and colleagues (1995) developed codes that would identify teaching for meaning. Rowan and associates (2002) constructed scales to measure cognitive demand. However, the actual outcomes were disappointingly slim. Perhaps differences are far more teacher than subject-matter based, as both of these studies would suggest. Or, perhaps the problem is with the measures. Rowan and colleagues (2002) were particularly forthcoming in what they saw as limitations to the *Prospects* data set. They noted that survey items were collected at the end of the year, teachers were instructed to answer across the entire year, and items were either answered across all subjects, or if they were subject specific, were coarse grained. They concluded that outcomes may underestimate the magnitude of any effects, including meaningful differences between mathematics and reading.

It is also possible that work so far has underestimated the challenges in crossing the borders across subject matters. To be sure, this work has laid an important foundation on which to build. But it may now be feasible to acknowledge the complexities and develop approaches for generalizing across disciplines while simultaneously maintaining subject matter integrity and identifying subject matter patterns that cut across individual teacher differences.

Challenges in Studying Teaching Across Two Disciplines

In this portion of the paper we briefly describe the study of mathematics and reading classrooms in which we were engaged, and highlight the challenges faced. The challenges include the tension between maintaining disciplinary specificity versus generalizing across disciplines, the differences in the language and foci in reading and mathematics education research, and the tradition of inequity of resources between mathematics and reading instruction.

Background

In this paper we draw on our work on and the data from a four-year study (2001-2005) of 4th and 5th grade classes in reading and mathematics. The purpose of the study was to learn more about what teachers and programs do to assist students who are struggling to acquire foundational skills in these subject areas. We selected 4th and 5th grades since this is a critical stage for students to acquire foundational literacy and numeracy skills (Valli & Croninger, 2001). Because of our particular interest in teacher success with low-achieving students, we located the study in one of the largest and most diverse school districts in the country, with 25 schools and approximately 70 teachers participated each year. We selected schools with moderate to high levels of poverty (30-90%) that had higher than expected levels of student achievement, that were nominated by knowledgeable insiders, or with whom we had professional ties.

Primary forms of data collection included classroom observations with time-sampling protocols and a high inference scale based on learner-centered principles (Alexander & Murphy, 1998), teacher daily logs of curriculum coverage, and interviews with principals, teachers, and educational specialists. As discussed later in the paper, standardized protocols were informed by, but not limited to, our conceptions of high-quality teaching derived from research on reading and mathematics instruction (Chambliss & Graeber, 2003), cognitive psychology, (Alexander & Murphy, 1998), and policy-practice relationships (Valli, Croninger, & Price, 2003). In this paper, we draw primarily but not entirely on one construct— *cognitively demanding instruction*—for illustrative purposes. The paper emphasizes perspectives and data from mathematics education but does this in the context of challenges of collecting data on the teaching of both mathematics and reading.

In carrying out this study we were interested from the outset in contributing to information about enduring questions within the two disciplines as well as producing some comparable data and attempting to provide a cohesive picture to teachers who are responsible for teaching both.

In many ways it seems to us that the primary challenges of working across subject matters are similar to the challenges faced by those attempting to teach in an interdisciplinary manner. A constant tension in the creation and delivery of interdisciplinary instruction is maintaining fidelity and depth of treatment to the disciplines involved while at the same time emphasizing their relatedness and intertwined use in everyday life. For example Grossman, Wineberg and Beers (2000) noted that in many attempts to integrate subjects "the disciplined part of disciplinary tends to fall away, leaving a body of information without the tools for evaluating its quality or warrant (p. 4). Similarly, Steen (1994) argued that integration can trivialize the differences between disciplines.

Discipline Specificity Versus Common Principles

This tension between the commonalities and differences was also a source of constant struggle for us as mathematics and reading educators, respectively. Whereas we wished to look at schooling from the perspective of the life of the teacher (most of whom in our study taught both mathematics and reading, even if not to the same group of children), we both shared beliefs in the notion that there is subject matter specific knowledge for teaching reading or mathematics. On the other hand, surely there are principles that would help guide a teacher pursuing effective instruction across subject matters. To what extent does the pursuit of general pedagogical principles undermine the pursuit of the unique aspects of effective teaching of reading or mathematics? This is a question we ponder, but have not answered; perhaps it is unanswerable.

The influence of discipline caused further tensions in our work in mathematics. Unlike reading, mathematics education has not only mathematics educators to listen to when attempting to grapple with what is effective teaching, but also mathematicians. As evidenced by the recent "Math Wars," not all mathematicians accept the portrait of effective teachers portrayed in documents such as the NCTM 2000 *Principles and Standards. for School Mathematics*. Indeed, this disagreement likely led to the position espoused in the National Research Council's sponsored and vetted publication, *Adding it Up* (Kilpatrick, Swafford, & Findell, 2001) that portrays four distinct portraits of high-quality instruction in mathematics. (Each vignette is accompanied with a discussion of the likely benefits and limitations of the portrayed approach.) Sorting notions of high quality teaching, like teaching itself, is highly complex.

The tensions in reading are as real, but arise more from a three-way clash among psychologists, sociologists, and scholars in English/language arts. Psychologists have identified reading processes that characterize expert reading and have explored effective instructional approaches for developing these processes (e.g., the National Reading Panel Report, 2000). Scholars who focus on English/language arts value instruction that fosters an aesthetic response to fine literature (e.g., Rosenblatt, 1978). Sociologists believe that all language skills develop in relationships with others and value establishing discourse communities to support the development of improved communication skills (Cohen, 1986; Swales, 1990). The most recent *Handbook of Reading Research* (Kamil, Mosenthal, Pearson, & Barr, 2000) presents all three perspectives with virtually contiguous chapters offering contrasting views of the goals of reading instruction and how to reach those goals.

In crafting the observation instrument and the teacher log for our own study, we began by reviewing research (both mathematics and reading) and looking at existing observation instruments. In mathematics we looked both at research supporting direct instruction models as well as those designed to reflect more cognitive or constructivist views. We listed attributes of direct instruction models and those of more constructivist models. Teacher and student activities favored in both of these approaches were incorporated in the instruments. In reading, we listed characteristics of reading instruction that reflected research from psychology, English/language arts, and sociology. Teacher and student activities, content, and text types that favored all three approaches were incorporated in the instruments.

One way in which we attempted to collect data that were related to a more general view of teaching was to consider ways in which ideas deemed important to the reading and mathematics educators were aligned with Alexander and Murphy's (1988) five learner-centered principles. This exercise helped us place some content specific concerns within a more generic framework. And the "Attribution Scale," our instrument for collecting high-inference data on the five learner-centered principles, is one way in which we will be able to observe and frame discussion of similarities and differences in teaching reading and mathematics.

Research Bases

In implementing integrated curricula, teacher knowledge is an issue not unrelated to the depth and of fidelity given to a discipline (e.g., Hammerness & Moffet, 2000; Mason, 1996; Steen 1994.) In our case, the parallel issue was depth of knowledge by researchers about the similarities and differences inherent in the disciplines and their respective research bases. We were fortunate in having faculty and graduate students who brought expertise in one of the areas (mathematics or reading) but also team members whose expertise in educational psychology helped us in our struggle with cross-disciplinary issues.

Structure of data collection instruments. This struggle to compare mathematics and reading instruction played out in a number of venues. First it was apparent in the construction of data collection instruments. Could we agree on categories of data we wanted to collect and what research bases we would use to help determine and define such categories? How similar did we want/could we have the options within such categories to be to still allow us both to make comparisons across subjects and yet feel that we would be able to speak to our individual content communities? Because classroom observers would collect data in both subject areas and teachers would code information for both reading and mathematics, there was some additional logistical press to keep data points, definitions, and instruments relatively similar. However, in looking at past studies within the two disciplines, we noted different emphases on characteristics of high quality teaching. Studies of the teaching of reading emphasize issues such as the relationship of phonics and comprehension, the degree of choice readers have in selecting text, and attention to various genre. Mathematics educators on the other hand attend to the role of basic skills versus problem solving, the cognitive demand of tasks presented to students, the attention paid to both the conceptual and procedural strands of mathematical proficiency and in particular the linking of the two, and attention to and discussion of multiple ways of completing problems or algorithms.

Definitions and language. A second challenge related to differences in research bases has to do with how disciplines use language or methods in similar but yet distinct ways. Wineburg and Grossman (2000) describe one school's history and English teachers working to create a community of learners among themselves prior to attempting curriculum integration for their high school. The authors describe the difficulties the teachers had in agreeing on approaches and in overcoming the notion that argumentation will lead to division. Wineburg and Grossman noted that "our suspicion is not that our group is exceptional or unrepresentative; rather our hunch is that other projects mask latent disciplinary differences and skirt conflict (p. 71). This same issue is discussed by Draper and Siebert (2004), two educators, one in mathematics education and one in literacy, who describe their journey in learning to communicate in meaningful ways.

Our own experience suggests terms used in one discipline may be rarely, if ever used in the other. For example, "cognitive demand" is an often used term in recent mathematics literature. Use a library search engine with "cognitive demand" and "mathematics" and you will find a fair number of entries. On the other hand, few entries appear when one uses "cognitive demand" and reading. The reverse is true when "community of discourse" is entered. However, as noted below, somewhat similar concepts appear in the literature for both subject areas.

In other instances similar terms may suggest very different concepts. Consider the idea of alternative answers or methods. In mathematics education, the notion of students presenting alternative methods, for solving a problem or accomplishing an operation in a non-standard way (e.g., solving 25 x 24, by thinking of (25 x 4) x 6) is almost always taken to be reflective of reasoning and higher order thinking. On the other hand alternative answers in reading to questions such as “How else do you think that story might have ended?” may result from personal experience or opinion and may draw on thinking that is less “higher order” or less dependent on formal knowledge.

Indeed Rosenshine (1979) described how different process-product researchers had coded questions of the ”alternate ending” type, noting that in some cases they were coded as higher level simply because more than one answer was possible. Such “higher level” questions, were then found to be negatively correlated with achievement. This lead Rosenshine to suggest that such questions might represent non-academic questions (p. 45). A reading educator might propose an alternative explanation: that encouraging students to consider such alternative answers could be poor test preparation for items with one and only one “correct” answer resulting in negative relationships between the items and achievement. The reading educator might also reject the idea that asking such questions in class is “nonacademic.” Indeed, such questions may be high order, even if they do not require technical knowledge in the same sense that alternative answers in mathematics seem to demand.

Negotiating differences in the meaning of terms across disciplines is further compounded by the way in which different researchers have defined terms even within a subject matter. For example, the *Prospects Study* ranked topics from whole numbers through algebra and noted that this could be “thought of as indexing the difficulty of the mathematics content covered (Rowan, Correnti, & Miller, 2002, p. 1548). This is very different from research that attempts to rate cognitive demand along more psychological process lines. For example, there are aspects of algebra that can be taught at through memorization and recall. On the other hand, there are aspects of number (e.g., compare two decimal of uneven length) that may be done by drawing on conceptual understanding of the relative size of place values to the right of the decimal or quite mechanically (e.g., to compare to decimals of unequal digit length, annex zeroes to the shorter until they are the same length. Compare as whole numbers). Analogously, in reading it could be possible to teach students to analyze spelling patterns according to the language origins of words or to memorize “rules” (*I* before *e* except after *c* and except in *weigh* and *neigh*.)

Differences in Resources. Another source of challenges centered on the differences in resources afforded to mathematics and reading and traditions in teaching the subjects. Time spent on reading is frequently double that of the time spent on mathematics (e.g., Knapp, et al., 1993, p. 76), more alternative resources are generally available in reading (e.g., both published basal reading series and trade books), and more reading specialists than mathematics specialists are employed at the elementary level. For example, in his study of Chicago area schools, Spillane (2005) noted that “across all schools in the study formally designated leadership positions for literacy outnumbered similar positions for mathematics three to one“(p. 388).

Not only does reading/language arts have more time devoted to it in the typical school day (Perie, Baker & Bobbitt, 1997, p. 27) a fact that was also true in our study, but what counts as “reading time” is likely much more complex than what counts as “mathematics time”. The exhortation to integrate reading and writing for example, makes it problematic to say that the time a student spends writing about what he or she just read is not “reading” time. In mathematics such distinctions and problematic situations rarely arise. In the end, we left general guidelines for the individual teacher to determine what 60-90 minute time period best captured his or her reading instruction, and observers planned their observations accordingly. Other evidence of time spent on reading can be drawn from the teacher’s daily log entries, but again, they identified the portions of the day spent on “reading.”

Consider also that reading frequently draws on texts representing different genres. During pilot work, we discovered that analyzing the specific genre that each student was using in the reading class proved difficult for our observers. And, yet, we considered genre to be an important feature of reading instruction. Consistently, researchers have found students in the middle grades to struggle more with exposition than with narrative text (e.g., Coté, Goldman, &

Saul, 1998, Kucan & Beck, 1996). Reading educators have speculated that children struggle because they receive less instructional experience with exposition than with narrative (Dreher, 2000). Thus, we wanted to be sure to capture all instruction with expository text. We taught observers how to distinguish the differences between exposition and narrative and left the finer distinctions (e.g., historical fiction, persuasive essay) up to the teachers' entries in their daily logs. This task was further complicated for observers because students in a class were often reading or writing different genres. In mathematics there is generally a textbook. Although teachers may draw on supplemental sources, these are frequently handouts for the entire class, artifacts relatively easy to collect.

In general the greater resources (both human, material, and time) spent in reading instruction, may make comparisons between the two more complex. Because of differences in absolute amounts of time for mathematics and reading, we report our observation results in proportion of total time. This choice may very well result in loss of comparability to some unknowable degree. For example, increased time to spend on a subject, may free teachers to feel they may allocate more time to higher order content. Greater differentiation of texts and the use of leveled texts may make alignment of material with an individual student's reading level somewhat easier than attempts at such alignment in mathematics. At any rate, children are spending more time on reading instruction than they are spending on mathematics, a difference that our use of proportions obscures.

Characteristics of High Quality Teaching Common to Math and Reading

In planning initial perspectives from which to view data, we looked at issues that crossed the two subject matter areas of reading and mathematics. The initial perspectives involved looking at data related to level of cognitive demand, the representation of the subject matter, and the existence of a community of discourse in the classroom. Each of these perspectives has received attention, including debate about the associated values, in the literature on effective teaching of mathematics. Similar, although not identical, perspectives exist within reading. These constructs are not unrelated, and we recognize the pitfalls of examining them separately. Nevertheless, we decided to begin our probing of the data by first looking at these individual clusters. For purposes of this paper, we explore the notion of cognitive demand in both reading and mathematics classrooms

Cognitive Demand

Few would argue that cognitive demand is unimportant to learning (although they might disagree about when higher cognitive demand should occur in an instructional sequence in mathematics or whether cognitive demand per se is a useful construct in reading). In mathematics education, Gamoran, Porter, Smithson, and White (1997) used a six-point scale to measure cognitive demand of specific content and found that such a measures (combining coverage and demand) was a better predictor of high school student gains on a problem solving test in mathematics than the more commonly used measure of content coverage alone.

More frequently, cognitive demand in mathematics instruction has been measured by examining “high order” and “low order” questions and tasks. Cognitive demand in the form of task complexity has been investigated in international studies of mathematics teaching at the eighth grade level (Hiebert, et al., 2003) and has some support in recent research as being important to middle school student’s achievement in mathematics. (Stein, Smith, Henningsen, & Silver, 2000).

High and low order questions and tasks have been defined in a number of ways (see for example, Doyle’s 1983 discussion of types of academic tasks). Many definitions of “high” and “low” level tasks or questions are in some way connected to Bloom’s (1956) taxonomy, with simple recall of facts or procedures labeled as lower and analyses and non-routine problem solving a task of higher order cognitive demand. Mathematics education on the other hand considers many forms of translation (especially across symbol systems) to be of higher order than Bloom’s rankings suggest. For mathematics classes we defined higher/lower order tasks and questions as shown in the excerpt from our observers' glossary (See Appendix.). This is probably closer to the definitions used by Stein, Smith, Henningsen, and Silver (2000) than a strictly Bloom's taxonomy based definition. The researchers in our project came to agree that we would consider cognitive demand to be reflected *by the degree to which the teacher presses for or evokes reasoning, reflection on learning, higher order thinking; or covers new/challenging content*. In this paper we look at teacher requests for higher order cognitive thinking, student responses that signal higher cognitive thinking, as well as the prevalence of content that demands conceptual, procedural, or linked conceptual and procedural proficiencies. This seems to align well, although not perfectly, with definitions used in the literature.

Even though the term "cognitive demand" rarely occurs in the reading literature, our definition certainly has both scholarly and research support as a feature of effective reading instruction. The *Standards for the English Language Arts (IRA & NCTE, 1996)* describes high quality teaching as encouraging student questioning, brainstorming, hypothesizing, reflecting and imaging. It also provides opportunities for students to invent new ways of using language by pursuing "imaginative risks" that depart from conventions and ways of expression (Chambliss & Graeber, 2003). Taylor, Pressley, and Pearson (2002) have summarized research in reading instruction that enhances student achievement. They concluded that the study by Knapp and associates (1995) demonstrated that effective reading instruction "emphasized higher order meaning making much more than lower-order skills" (Taylor, Pressley, & Pearson 2002, p. 364). Taylor and colleagues (2002) concluded that five studies of effective instruction, including the work by Knapp and associates (1995), Taylor, Pearson, Clark, and Wlapole (2002), and the original *Prospects* report, could be characterized as fostering higher order thinking. So, although the curriculum framework for the English/Language Arts and a recent review of effective reading instruction did not use the term "cognitive demand," we felt confident that applying our definition to reading instruction was justified. However, because the preponderance of work on this construct has been conducted for mathematics instruction, our discussion below focuses far more on mathematics instruction than on reading.

Teacher Questions and Student Oral Responses

Researchers' attention to the level of teachers' questions and student extended responses has been quite prevalent in the last 30 or so years of literature in mathematics education. Teachers high and low order questions as well as students simple or extended responses are variables that were explored in the process-product studies of primary grades (e.g., Rosenshine & Stevens, 1984), fourth grade (Good, Grouws & Ebemeir, 1983), and seventh and eighth grade (Evertson, Emmer, Brophy, 1980). Most of these explorations of direct instruction and active teaching models found benefits to standardized text scores in classrooms in which teachers posed primarily low-level questions that required students to give simple answers. For example, Rosenshine (1979), described direct instruction as referring to "teaching activities where goals are clear to students, … questions are at a low cognitive level so that students can produce many correct responses and feedback to students is immediate and academically oriented" (p. 38). Similarly Good and Grouws (1977) found that in whole class settings, more effective teachers asked fewer process questions ("questions demanding integration of facts, explanation"), than did less effective teachers (p. 52). On the other hand, Evertson, Emmer, and Brophy (1980) found that of nine teachers using a direct instruction approach, the more effective teachers asked more questions of all types; however, process questions represented a greater proportion of the more effective teacher's questions than they did for less effective teachers (p.172). Further, in 1981, Redfield and Rousseau challenged the findings concerning low level questions. They reported that their meta-analysis of studies, including training experiments and skill experiments involving grade levels and subjects ranging from high school chemistry courses to fourth grade reading, found that "gains in achievement can be expected when higher cognitive questions assume a predominant role during classroom instruction" (p. 237). Stodolsky (1988) also reported the highest level of student involvement (essentially engagement) in fifth grade mathematics when the classroom segments included application or higher mental processes (pp. 83-84).

The findings suggesting the effectiveness of a preponderance of low level questions have been much debated in the literature on several grounds (e.g., Peterson, 1988). One is that the standardized tests used to assess achievement in these process-product studies primarily tested lower level skills and gave little attention to students' understanding or processes. A second criticism of the process-product approach is offered by Peterson (1988) who cited Bennett's 1986 argument that higher order skills may be those necessary for success in learning more mathematics. And, since achievement deficits are likely to be magnified in secondary school, it is extremely important that higher order thinking become a part of elementary mathematics instruction. Recent proponents of teaching for understanding also suggest that students must have opportunities to "articulate what they know" (Carpenter & Lehrer, 1999, p. 24) and "students expect that their teacher and their peers will want explanations as to why their conjectures and conclusions make sense and why a procedure they have used is valid for the given problem" (p. 26). This suggests a need for high order questioning that calls for explanation, justification, alternative answers and reasoning.

One example of the impact of high level questions in reading instruction is found in the heavily researched strategy of reciprocal teaching (see the National Reading Panel Report, 2000). Although the eventual goal in reciprocal teaching is for students to take over the role of the teacher, initially teachers lead student discussions by asking students to question the text, summarize, predict what will come next, and ask for clarification when they do not understand. More recently, Beck and colleagues (Beck, McKeown, Hamilton, & Kucan, 1997) have developed an effective instructional approach called *Questioning the Author* in which the teacher asks students questions about the choices an author has made (particularly focusing on inconsistencies and unclear writing) with the goal to understand what the author means. Students are expected to explain their thinking.

Student Tasks

While attention to questioning in mathematics classes has been studied for some time, Doyle (1983) noted that attention to the nature of the specific tasks or work students are asked to do is a relatively recent phenomena. Indeed in the late 1980s and early 1990s the importance of tasks used in teaching was suggested as important to learning in a number of policy documents. For example, the National Council of Teachers of Mathematics' (1991) *Professional Standards for Teaching Mathematics* emphasized the importance of "worthwhile mathematical tasks in instruction, listing it as the first standard in that publication. Worthwhile tasks were described as those that (among other things) "call for problem formulation, problem solving, and mathematical reasoning; promote communication about mathematics; display sensitivity to, and draw on students' diverse background experiences and dispositions" (p. 15).

Renkl and Helmke (1992) suggested that the conflicting results obtained concerning the impact of high and low level questions, might be explained by classifying questions and tasks differently and by assessing finer effects by looking at different aspects of achievement. They note that many studies classified all process questions as high-level; and they argue, in a fashion similar to us, that not all questions about process are high-level. They argue that process questions and tasks should be classified as to whether whether they draw on mechanical knowledge or meaning. Renkl and Helmke defined "performance-oriented tasks" as those requiring mechanical knowledge (i.e., single facts or procedures learned by rote) and "structure-oriented" tasks as those requiring meaningful knowledge (principles, concepts and conceptually based procedures (p. 48). Their study of 33 third-grade mathematics classes suggested that "performance oriented tasks primarily promote the automatization of basic arithmetic skills, whereas structure oriented tasks foster mathematical problem solving skills" (p. 52). This gives support for our coding guides for procedural questions and tasks in mathematics. If the procedures learned were new, requiring application of meaning, they were coded as high-level tasks. If the procedures were routinized, the questions and tasks were coded as routine.

A number of researchers have built on Doyle's (1983) work on the nature of academic tasks and students' means of circumventing task demands. In mathematics education, Hiebert and Wearne (1993) study of six second-grade mathematics classrooms led them to suggest that "certain kinds of instructional tasks and discourse encourage more productive ways of thinking' (p. 421). Students whose instruction involved more problem solving and reflection as compared to more procedural instruction did fewer problems in class but at the end of the year were more successful on novel problems.

A project that has become well known for its attention to students' work on mathematical tasks as well as what supports and what undermines the level of tasks as they are implemented and carried out, is the QUASAR Project (Quantitative Understanding: Amplifying Student Achievement and Reasoning). The project created supports for reform type mathematics instruction in urban, high-minority, middle schools and documented instruction and measured student outcomes. Silver and Stein (1996) reported on finding that

> …learning gains were especially positive in classrooms that could be characterized by the set up and implementation of instructional tasks that encouraged high-level thinking and reasoning and the use of multiple solutions strategies, multiple connected representations, and mathematics explanations. By contrast, in those classrooms where instruction tended to be procedural in nature and to require only single solution strategies, single representations, and little or no mathematical communication, student learning gains were very small (p. 506).

Similar claims exist about the importance of high-level tasks in reading. In their review of instructional studies in reading, Taylor, Pressley, and Pearson (2002) focused on the types of tasks that were associated with student achievement in reading. Across studies, they concluded that students in classes with excellent elementary reading teachers (as measured by student achievement or reputation) are engaged much of the time in reading and writing tasks that are based on the very strategies on which they need to work. The goal of these tasks is meaning-making rather than the practice of discrete skills. Thus, students complete tasks that will help them understand what they are reading. Any skill tasks are in service of that end.

Demand of the Content

Content, when thought of in terms of mathematics, signals to most a list of topics, such as numeration, multiplication, area, perimeter, arithmetic mean, etc. But almost any content can be taught at various levels of sophistication and abstraction. We also believed that teachers would likely be reliable sources of information about what content they had taught on a given day. Thus, for our study we elected to have teachers record the topic(s) taught from a pull down menu in their daily log, and utilized the opportunity afforded in observations to assess whether the content was addressed procedurally, conceptually or involved the linking of the two. The roles of conceptual knowledge, procedural fluency[1] and the linking of the two have been of enduring interest and contention in mathematics education (Hiebert & Lefevre, 1986, p. 1). Because linking demands explanation and justification as well as connecting a procedure with underlying concepts or connecting across representations, the mathematics educator felt that a measure of time spent on linking would be a conservative measure for time when content was cognitively demanding. Certainly some classroom episodes coded conceptual and even some coded procedural may be cognitively demanding for some or all students. But whether the student response to the request, “Define a trapezoid,” is demanding or merely elicits a memorized response is unclear to an observer. Similarly as students work on procedures such as renaming fractions, this may or may not be cognitively demanding for a student and in a classroom situation is not a distinction an observer can make. Thus, in mathematics, we did not make claims about how high or low level a task was based solely on the topic at hand. We relied

[1] It should be noted that procedures need not involve arithmetic. For example, constructing a circle graph, measuring an angle, or drawing a sketch to decide how to tackle a word problem all involve procedures.

on teacher logs to capture content in terms of topics and used the categories of conceptual, procedural, and linking to code content during observations of mathematics classes.

The clearest explication of content that would support cognitive demand in reading is in the English/Language Arts framework, which actually focuses more on strategies than content per se. The standards were based on the premise that all learners learn language by using it purposively to communicate with others. A central goal of English/Language Arts instruction must be to enable learners to use language to address their own needs and the needs of their own families, communities, and society as a whole. Because students develop language in different ways and at different rates, adaptability would be the hallmark of high quality teaching as teachers respond to the learning needs that arise. If teachers should adapt content choices to the needs of the students in a class, then the content would be cognitively demanding. The needs of the learner were to be at the core of all instructional decisions including content.

Nonetheless, the Framework explains, in order to learn to communicate well, students would need to be exposed to a wide range of print and nonprint text that represented fiction and nonfiction genres, classic and contemporary works, different difficulty levels, and diverse cultures. It would introduce students to and give them practice in an array of processes and strategies for comprehending and producing texts, including the use of background knowledge to construct meaning, effective strategies for fluently identifying words, study strategies to enhance learning and recall, and systematic processes for approaching writing. It would include the study of the systems and structures of language and of language conventions, including grammar, punctuation, and spelling (and how these can vary by context). The exact cognitive demand of each of these topics would depend on whether they specifically built on student need.

A Comparison of High Cognitive Demand in Reading and Mathematics Classes

In the larger study, we are examining change in teaching practices over time and how those changes might be related to changes in policy. For example, Valli, Croninger, and Buese

(2006) examined the decline of cognitive demand over the three primary years of data collection. In this paper, we were interested in how cognitive demand compares in math and reading and on the challenges of making such comparisons. We also wanted to know whether the same teachers who were cognitively demanding in one subject were cognitively demanding in the other subject. For these reasons, we decided to report on data from one year only. Specifically, we chose Year 3, the year with the most observed lessons. There were 606 observed lessons for math and 552 for reading. These lessons involved 74 math teachers and 71 reading teachers. Of these, 69 teachers taught both subjects.

In keeping with the cognitive demand construct, we decided to examine three important aspects of teaching: teacher activity, student activity, and content. The codes listed below were chosen (after much discussion among the experts in our study) from the observation instruments to represent high cognitive demand for these three aspects of teaching.

Teacher Activity, Student Activity, and Content

For math, teacher activities considered to be cognitively demanding included requests for: student reflection on learning, alternative methods or strategies, student self assessment, elaboration of a student response, and attention to a student's response. It also included posing or elaborating on problem, task, or high order question. For reading, the codes were nearly identical. The only difference was that, for reading, requests for alternative answers replaced requests for alternative methods or strategies.

Student activities that represented high cognitive demand were also defined similarly for math and reading. For math, they included students responding with: conjectures, explanations or justifications, or alternative methods. They also included students working on a problem, task, or extended writing in mathematics. For reading, high cognitive demand was defined as students responding with: hypotheses or prediction, explanations or justifications, alternative answers, or elaborated answers. It also included when students did extended writing or created a graphic organizer or an outline. Performing speeches or plays was also considered cognitively demanding.

Cognitively demanding content was defined quite differently for mathematics and reading. For mathematics, only when procedures were linked to their related concepts was the "conent" deemed to be cognitively demanding. For reading, cognitively demanding content included aspects of both reading and writing. Content that focused on a text's theme or main idea or genre, or personal responses to the text would be cognitively demanding whether for reading or writing. Because students have so much less experience with exposition and poetry than narrative, considering the design of an exposition or the elements in a poem would create high cognitive demand for both modes, whereas the same type of content for narrative would not. Even content that covered writing that was unrelated to reading would have high cognitive demand if it concerned prewriting, composing, editing, or revising. In order to cover these different types of content, the observational protocol for reading had many more items under Content than the protocol for mathematics.

Findings

The proportion of time that teacher activity was cognitively demanding was similar for math and reading. For math, the mean was .11 (*SD*=.10). For reading, the mean was .10 (*SD*=.10). The proportion of time that student activity was cognitively demanding was also similar for math and reading. For math, the mean was .17 (*SD*=.14). For reading, it was .18 (*SD*=.13). See Table 1 for means and standard deviations. These outcomes suggest similarities between the teaching in the two subject areas.

In contrast, apparent differences were found for high cognitive content in math and reading. For math, the mean proportion of time the content was cognitively demanding was .13 (*SD*=.16). For reading, it was .29 (*SD*=.24). Whereas high cognitive teacher and student activity can be viewed in similar ways across reading and math, the fundamental differences in the subject areas precluded a common definition for cognitively demanding content. Judging solely from the data, it would seem that more cognitively demanding content is occurring in reading classes. But this conclusion must be tempered by the fact that these constructs were defined differently for mathematics and reading. Interpreting his finding is difficult due to the challenges in studying mathematics and reading instruction. This finding speaks directly to the challenges discussed throughout this paper.

In summary, teachers were asking students to engage in cognitively demanding activities roughly 10% of the time during both reading and math classes. Also for reading and math, students were working on cognitively demanding tasks or providing cognitively demanding responses slightly less than 20% of the time. In other words, the occurrence of these activities was relatively low for both subjects. On the other hand, students seemed to be engaged in cognitively demanding content more than twice as often in reading than in mathematics. This may be a result of how “cognitively demanding” content was defined in the two subject areas. As we implied above, cognitively demanding content in mathematics tends to be related to a specific formal knowledge, whereas cognitively demanding content in reading may be related to personal as well as formal knowledge. Thus range of tasks and the range of ways in which tasks can be responded to in cognitively demanding ways may be much greater in reading than in mathematics. The increase in items may have resulted in a more fine-tuned focus on content in reading than in mathematics, and therefore the difference that we obtained may be an artifact of our measurement instrument. Or, this difference may be a real distinction in how cognitively demanding the content was in the two subject areas. Although our data are reported as proportions, a greater emphasis on higher order content may be facilitated by the greater allocation of time to reading than to mathematics.

Table1

Proportion of Class Time Spent on High Cognitive Demand

	Teacher Activity	Student Activity	Content
Mathematics			
M	.11	.17	.13
N	606	606	606
SD	.10	.14	.16
Reading			
M	.10	.18	.29
N	552	552	552
SD	.10	.13	.24

Although the level of cognitive demand in general seems to be similar across subjects for teacher activity and student activity, it does not follow that the level of cognitive demand is similar across subjects for the same teacher. In fact, there was evidence that teachers whose classrooms were cognitively demanding in math were not always the ones whose classrooms were cognitively demanding in reading. Of the 69 teachers who taught both math and reading, only 19 (28%) were above the mean for cognitively demanding teacher activity in both math and reading. For high cognitive student activity, 12 of the teachers (17%) scored above the mean for both math and reading. For content, eight of the teachers (12%) scored above the mean for both subjects. In general, these results are consistent with those found by Knapp, Shields, and Turnbull (1992). They found that only 15% of the teachers in their study adopted alternative practices in more than two subject areas (p.10).

Conclusions and Implications

In our efforts to explore the challenges faced in seeking to characterize and compare high quality teaching for reading and mathematics, we discussed some of the challenges faced in our and others' study of teaching across these subjects, and used the perspective of cognitive demand to illustrate some specifics of such work. Below we summarize our discussions and suggest implications for future work.

Observations about Research

As noted at the beginning of this paper, older process-product studies often observed mathematics and reading instruction with the same or similar instruments, but often reported data devoid of any distinctions between the two subject matters. Somewhat later studies of that era included one subject matter or the other, not both (cites). Rowan et al. studied both and highlighted some differences, but they did not attempt any explanations of these differences. On the other hand Knapp et al. (1993), who studied both reading and mathematics, collected different data on the two subjects but did bring the results together under the more abstract notion of "teaching for meaning" which they characterize across subjects to some extent and within subjects to a greater extent.

Clearly there are substantial challenges to looking across subjects and maintaining the integrity of the subject matters. The differences in the epistemology of the two subjects raises a number of issues in any attempt to find similarities across the teaching of the two (e.g., Alexander, Chambliss, & Price, 2006). It is perhaps not surprising, then, that the research bases on teaching of the two subject matters differ in the enduring questions addressed. Further, while activities may appear to be similar across the subject matters, they may indeed be very different. And this is further compounded by the way in which different researchers have defined terms even within a subject matter. We noted, for example, the differences in proxies used for cognitive demand, both within subjects and across subjects.

Our Work

First, we found commonalities across reading and mathematics. The amount of cognitive demand for teacher activity and student activity was very close for both subject areas across this relatively large sample and data collection over an entire school year. We might wish the proportions were higher, but we would wish that for both subject areas.

Second, we also were able to measure the amount of cognitive demand for the content covered within each subject area. Of the two subject areas, mathematics seems the more coherent. The proportions of cognitive demand for teacher activity, student activity, and content ranged between .11 and .17. Presumably, teachers were requesting higher order responses and posing or elaborating on a higher order problem, task or question, and students were making higher order responses or working on higher order tasks involving higher order content. In other words, teachers and students were linking conceptual understanding with procedural knowledge. Reading is less coherent. The cognitive demand of teacher activity, student activity, and content ranged from .10 to .29. The outlier was content. If we imagine the relationship in reading, there would be instances where teachers and students were involved in activities that had little cognitive demand, but the potential cognitive demand of the content was high. For example, the teacher could have been watching students complete a worksheet of multiple-choice or short-answer questions about the major ideas in an expository text about Benjamin Franklin. Neither the teacher activity or the student activity would be high cognitive demand, but the content would be.

Third, as we have already reported, only a small proportion of teachers who taught both mathematics and reading demonstrated higher than average cognitive demand in both subject areas. We are not quite sure what to make of it at this time, but the proportions differed for teacher activity, student activity, and content for these teachers. Over a quarter of them were above the mean for cognitively demanding teacher activity in both subject areas. However, only 17% scored above the mean for cognitively demanding student activity, and a mere 12% scored above the mean for cognitively demanding content. Taylor, Pearson and colleagues (2002) identified what they termed “preferred interaction style” for the reading teachers in their study. The teachers in our study who were more likely to press for higher order cognition than the rest of their cohort across mathematics and reading may have had a preferred instruction style that they were able to generalize across content areas. However, fewer teachers were able to generalize their style across student tasks and even fewer across two types of content. Perhaps these last two findings are reflective of the differences in teaching mathematics and reading.

Developing observational and teacher log protocols that could capture the common instructional possibilities in both reading and mathematics, but that also could capture content differences between the two subject areas, allowed us to compare and contrast. We have described both similarities and differences in the cognitive demand of the two subject areas. In out concluding section, we draw implications from these findings.

Implications for Teacher Education

As we begin to relate our classroom measures with achievement we hope that the care we took in defining the variables for our instruments in each of the subject matter areas and our in particular our own understanding of the locus of the difficulties in comparing them, may help us in reporting in ways that are useful to researchers as well as teachers.

Further, by looking at such carefully defined measures consistently across time (as reported in Valli, Croninger, & Buese, 2006) we believe we have some evidence that some policy changes, e.g., No Child Left Behind, do reach the classroom level. The state in which we work has substituted a test that is primarily although not exclusively a multiple choice test (to facilitate high reliability and individual scores) for a prior matrix sampling test that involved more performance tasks and more writing. Teaching to two such different tests may indeed require and result in different strategies for high achievement at least fin the short-term. It does appear that teachers are shifting focus to some extent. Tracking the impact of policy implementation should be carried out if we are to learn from experience. Whether or not such changes are in the public good in the long run is a question of what values we wish to reflect in schools.

Current pressures for high-quality teaching (when defined as teaching that raises achievement test scores) have made cross-subject comparisons more necessary than ever. Regardless of how “high-quality teaching is defined, some characteristics of teaching are inherently domain-specific, others may be applicable to multiple domains. If so, implications exist for preservice teacher education. Those learning to teach could benefit from researchers’ efforts to cross the heavily guarded borders of subject domains. If there are indeed some teaching practices considered to be high-quality in both math and reading, a clearer understanding of them can lead to more coherent education programs, particularly for elementary education majors. The present paper has attempted to share the challenges and early findings of such an undertaking.

For those similarities that exist only at relatively abstract levels, how might they be conveyed to and recognized by teachers? What organization of examples from different subjects might be valuable in helping teachers understand the similarities? Or are the similarities and subtle differences in meaning of terms so overwhelming that the periodic calls for elementary specialists in subject matter teaching should be given greater heed.

Reference List

Alexander, P.A., Chambliss, M., & Price, J. (2006, April). Ontological and epistemological threads in the fabric of pedagogical research. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Alexander, P. A., & Murphy, P. K. (1998). The research base for APA's learner-centered psychological principles. In N.M. Lambert & B.L. McCombs (Eds.), *Issues in school reform: A sampler of psychological perspectives on learner-centered schools* (pp.25-66). Washington DC: The American Psychological Association.

Beck, I. L., McKeown, M. G., Hamilton, R., & Kucan, L. (1997). *Questioning the author: An approach for enhancing student engagement with text.* Newark, DE: International Reading Association.

Bloom, B. (Ed.). (1956). *Taxonomy of educational objectives: The classification of educational goals* (Handbook I: Cognitive domain), New York: David McKay.

Brophy, J.E., & Good, T.L. (1986). Teacher behavior and student achievement. In M. Wittrock, (Ed.), *Handbook of research on teaching (3rd edition)* (pp. 328-375). New York: MacMillan

Carpenter, T., & Lehrer, R. (1999). Teaching and learning mathematics with understanding. In E. Fennema & T.A. Romberg (Eds.), *Mathematics classrooms that promote understanding* (pp. 19-32). Mahwah: NJ: Erlbaum.

Chambliss, M., & Graeber, A. O. (2003, April). Does Subject Matter *Matter*? Paper presented at the annual meeting of the American Educational Research Association, Chicago.

Cohen, E. (1986). *Designing groupwork: Strategies for the heterogeneous classroom.* New York: Teachers College Press.

Coté, N., Goldman, S. R., & Saul, E. U. (1998). Students making sense of informational text: Relations between processing and representation. *Discourse Processes, 25,* 1-53.

Croninger, R., Valli, L. & Price, J. (2003, April), Mapping the policy environment for high-quality teaching: Can we get there from here? Paper presented at the annual meeting of the American Educational Research Association, Chicago.

Davidson, D. M., Miller, K.W., & Metheny, D. L. (1995). What does integration of science and mathematics really mean? *School Science and Mathematics, 95*(5), 226-230.

Doyle, W. (1983). Academic work. *Review of Educational Research, 53,* 159-199.

Draper, R.J., & Siebert, D. (2004). Different goals, similar practices: Making sense of the mathematics and literacy instruction in a *Standards*-based mathematics classroom. *American Educational Research Journal, 41*, 927-962.

Dreher, M. J. (2000). Fostering reading for learning. In L. Baker, M. J. Dreher & J. T. Guthrie (Eds.), *Engaging young readers: Promoting achievement and motivation* (pp. 68-93). New York: Guilford Press.

Erickson, F. (1986). Qualitative methods in research on teaching. In M. Wittrock (Ed.), *Handbook of research on teaching (3rd edition)* (pp. 119-161). New York: MacMillan.

Evertson, C., Anderson, C., Anderson. L., & Brophy, J. (1980). Relationships between classroom behaviors and student outcomes in junior high mathematics and English classes. *American Educational Research Journal*, *17*(1), 43-60.

Evertson, C.M., Emmer, E. T.,& Brophy, J. E. (1980). Predictors of effective teaching in junior high mathematics classrooms. *Journal for Research in Mathematics Education, 11*, 167-178.

Gamoran, A., Porter, A., Smithson, J, & White, P. (1997). Upgrading high school mathematics instruction: Improving learning opportunities for low-achieving, low-income youth. *Educational Evaluation and Policy Analysis 19*, 325-338.

Good, T. L., & Brophy, J.E. (1984). *Looking in classrooms* (Third edition). Cambridge, MA: Harper & Row.

Good, T.L., & Grouws, D. A. (1977). Teaching effects: A process-product study in fourth grade mathematics classrooms. *Journal of Teacher Education 28,* (3) 49-54.

Good, T, L., Grouws, D. A., & Ebemeir, H. (1983). *Active mathematics teaching*, New York: Longman.

Grossman, P., Wineberg, S., & Beers, S. (2000). Introduction: When theory meets practice in the world of school. In S. Wineburg & P. Grossman (Eds.), *Interdisciplinary curriculum challenges to implementation* (pp.1-16). New York: Teachers College Press.

Hammerness, K., & Moffet,K, 2000. The subjects of debate: Teachers' clashing and overlapping beliefs about subject matter during a whole-school reform. In S. Wineburg & P. Grossman (Eds.), *Interdisciplinary curriculum challenges to implementation* (pp.134-152). New York: Teachers College Press.

Hiebert, J., & Lefevere, P. (1986). Conceptual and procedural knowledge in mathematics: An introductory analysis. In J. Hiebert (Ed.) *Conceptual and procedural knowledge: The case of mathematics* (pp.1-27). Hlllsdale, NJ: Erlbaum.

Hiebert, J., & Wearne, D. (1993). Instructional tasks, classroom discourse, and students' learning in second-grade mathematics. *American Educational Research Journal, 30*, 393-425.

IRA and NCTE (1996). *Standards for the English language arts*. Published by the International Reading Association and the National Council of Teachers of English.

Kamil, M. L., Mosenthal, P. B., Pearson, P. D., & Barr, R. (2000). *Handbook of reading research: Volume III.* Mahwah, NJ: Erlbaum.

Kilpatrick, J., Swafford, J., & Findell, B. (Eds.). (2001). *Adding it up: Helping children learn mathematics*. Washington, DC: National Academy Press.

Knapp, M.S., & Marder, C. (1992). *Study of academic instruction for the children of poverty: Volume 2 study design and technical notes*. Prepared by SRI International and Policy Study Associates. Washington, DC: Office of Policy and Planning, U.S. Department of Education.

Knapp, M.S., Shields, P, M., & Turnbull, B.J. (1992). *Academic challenge for the children of poverty: Summary report*. Prepared by SRI International and Policy Study Associates . Washington, DC: U.S. Department of Education.

Knapp, M.S., Adelman, N.C., Marder, C., McCollum, H., Needels, M., Shields, P., et al. (1993). *Academic challenge for the children of poverty:* Volume *I Findings and conclusions.* Prepared by SRI International and Policy Study Associates . Washington, DC: Office of Policy and Planning, U.S. Department of Education. (ERIC Document #ED 358213).

Knapp, M.S., Shields, P. M., & Turnbull, B. J. (1995). Academic challenge in high-poverty classrooms. *Phi Delta Kappan, 76, 770-777.*

Kucan, L., & Beck, I. L. (1996). Four fourth graders thinking aloud: An investigation of genre effects. *Journal of Literacy Research, 28,* 259-287.

Lehman, J. R. (1994). Integrating science and mathematics: Perceptions of preservice and practicing elementary teachers. *School Science and Mathematics, 94,* 58-64.

Mason, T. C. (1996). Integrated curricula: Potential and problems. *Journal of Teacher Education, 47*(4), 263-270.

Medley, D. (1979). The effectiveness of teachers. In P. L. Peterson and H. J. Walberg (Eds.) *Research on teaching: Concepts, findings and implications* (pp. 11-27). Berkeley, CA; McCutchan.

National Council of Teachers of Mathematics. (1991). *Professional standards for teaching mathematics*. Curriculum and evaluation standards for school mathematics. Reston, VA: National Council of Teachers of Mathematics.

National Council of Teachers of Mathematics. (2000*). Principles and standards for school mathematics*. Reston, VA: National Council of Teachers of Mathematics.

National Reading Panel (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction:*

Reports of the subgroups. Washington, DC: National Institute of Child Health and Development.

Peterson, P.L. (1988). Teaching for higher-order thinking in mathematics: The challenge of the next decade. In D. Grouws and T.J. Cooney (Eds.),*Perspectives on effective mathematics teaching* (pp. 2-26). Reston, VA: National Council of Teachers of Mathematics.

Perie, M., Baker, D. P., & Bobbitt, S. (1997). Time spent in teaching core academic subjects in elementary schools: Comparisons across community, school, teacher, and student characteristics. Washington, DC: National Center for Educational Statistics.

Redfield, D., & Rousseau, E. (1981). A meta-analysis of experimental research on teacher questioning behavior. *Review of Educational Research*, 51, 237-245.

Renkl, A., & Helmke, A. (1992). Discriminant effects of performance-oriented and structure-oriented mathematics tasks on achievement growth. *Contemporary Educational Psychology 17*, 47-55.

Rosenblatt, L. M. (1978). *The reader, the text, the poem: The transactional theory of the literary work.* Carbondale: Southern Illinois University Press.

Rosenshine, B.V. (1979). Content, time and direct instruction, In P. L. Peterson and H. J. Walberg (Eds.) *Research on teaching: Concepts, findings and implications*, (pp. 28-56). Berkeley, CA; McCutchan.

Rosenshine, B.V., & Stevens, R. (1984). Classroom instruction in reading. In P.D. Pearson (Ed.), *Recent research in reading*. New York: Longman.

Rowan, B., Correnti, R., & Miller, R. J. (2002). What large-scale, survey research tells us about teacher effects on student achievement: Insights from the *Prospects* study of elementary schools. *Teachers College Record, 104*, 1525-1567.

Silver, E.A., & Stein, M.K . (1996). The “revolution of the possible” in mathematics instructional reform in urban middle schools.” *Urban Education, 30*, 476-521.

Spillane. J.P. (2005). Primary school leadership: How the subject matters. *School Leadership and Management. 25(4), 383-397.*

Steen, L. A. (1994). Integrating school science and mathematics: Fad or folly? In D. F. Berlin (Ed.), *NSF/SSMA wingspread conference: A network for integrated science and mathematics teaching and learning.* (pp. 7-12). Bloomsberg, PA: School Science and Mathematics Association.

Stein, M.K.,.Smith, M.S. S., Henningsen, M.A., & Silver, E. A. (2000). Implementing standards-based mathematics instruction: A casebook for professional development. New York: Teachers College Press.

Stevens, R., Wineburg, S, Herrenhkohl, L.R., & Bell, P; (2005). Comparative understanding of school subjects" past, present, and future. *Review of Educational Research, 75*, 125-157

Stodolsky, S. S. (1988). *The subject matters: Classroom activity in math and social studies.* Chicago: University of Chicago Press

Swales, J. (1990). *Genre Analysis : English in Academic and Research Settings***.** Cambridge, UK: Cambridge University Press.

Taylor, B. M., Pearson, P. D., Clark, K., & Walpole, S. (2002). Effective schools and accomplished teachers: Lessons about primary-grade reading instruction in low-income schools. In B. M. Taylor & P. D. Pearson (Eds.), *Teaching reading: Effective schools, accomplished teachers* (pp. 3-72). Mahwah, NJ: Erlbaum.

Taylor, B. M., Pressley, M., & Pearson, P. D. (2002). Research-supported characteristics of teachers and schools that promote reading achievement. In B. M. Taylor & P. D. Pearson (Eds.), *Teaching reading: Effective schools, accomplished teachers* (pp. 361-373). Mahwah, NJ: Erlbaum.

Valli, L., & Croninger, R. (2001). High Quality Teaching of Foundational Skills in Mathematics & Reading. Proposal (funded) submitted to the National Science Foundation's Educational Initiative Program. College of Education, University of Maryland.

Valli, L., Croninger, R., & Buese, D. (2006, April). Studying High-Quality Teaching in a Highly-Charged Policy Environment. Paper presented at the annual meeting of the American Educational Research Association, San Francisco: CA.

Wineburg, S., & Grossman. (2000). Scenes from a courtship: Some theoretical and practical implications of interdisciplinary humanities curricula in the comprehensive high school. In S, Wineburg & P. Grossman (Eds.), *Interdisciplinary curriculum challenges to implementation* (pp.57-73). New York: Teachers College Press.

Appendix: Selected Terms as Defined for the Mathematics Observation Instrument in the Classroom Coding Manual Glossary

Problem/task/higher order question: A task or exercise is only a problem if the teacher's action indicates that the student has not learned and is not presently learning a routine to apply. Students not taught, or who have not grasped, an algorithm for renaming a mixed numeral as a fraction, may find this task to be a problem. However, if they have mastered or are learning the routine--divide the denominator into the numerator-- this is not a problem. Higher order questions require more than simple recall. Requests for why something is true or why a procedure works are generally higher order questions. If it is really arguable whether a task or question is higher or lower order, classify it as higher order.

Routine exercise or low order question: A routine algorithm or procedure that the student seems to have already learned or is in the process of learning. A question that asks for a simple answer, a simple characteristic, how to do a routine algorithm, recall of a definition are low order questions.

Linking Procedural and Conceptual - The teacher or the task requires students to make explicit connections between a procedure and a concept (see definitions below). The question or task or statement links a *why* or *what* explanation to a *how* explanation.

> **Linking questions** often ask for the justification of a procedure or procedural rule. They ask students to tell why rules or procedures work. *"Janice, when you round a number like 67 to the nearest ten, why is it that you look at the digit in the one's place?"* Or *"What does the fact that 7 is greater than 5 have to do with how you round? "*
>
> Linking tasks: **Students are representing common fractions with fraction circles and joining two such representations to illustrate addition. They then write (or orally explain) the symbolic sentence for that addition example. They are asked to use words to relate the symbolic statement to the manipulative.**
>
> **Linking responses:** Example, *"3/4 is the same as 30/40 because when you multiply the numerator and denominator both by ten, it is like multiplying 3/4 by 10/10 or 1, and the product of 1 and any number is just that same number."*

Conceptual -- has to do with understanding *why* or *what.* Conceptual knowledge is considered as composed of ideas in relationship to one another, related knowledge,

> **Conceptual questions** are often but not always why or what questions. For example:
> *Why do you say that 7/8 is a number close to 1?*
> *What's the definition of the greatest common factor of two numbers?*
> ***When we round to the nearest ten, why do those numbers whose one's digits are 6, 7, 8 or get rounded up?***
> *What relationship does* ***π*** *symbolize?*
> *Will this whole liter (pointing at liter container) fit in this coke can (pointing at normal size soda can)?*
> *How do you know the angle will measure more than 90 degrees?*
>
> **Definitions are conceptual**: Some low order questions are conceptual. For example, asking for the definition of a rectangle is conceptual. Or, asking students to list all the

measuring tools they can think of is conceptual, even if it is simple review, because the focus is on the concept of a measuring tool not on the procedure of using the tool. If it is really arguable whether the focus is conceptual or procedural, code as conceptual.

Conceptual response: Gives an underlying reason or rationale, relates information to basic understandings – "*Rounding to the nearest ten is naming the nearest ten. If you take some space on the number line, say 20-30, then 26, 27, 28 and 29 are closer to 30 than to 20 -- so we round up.*"

Procedural - involves knowing *how to*. It involves applying a procedure, an algorithm, a routine, or how to write or read a symbol. If it is really arguable whether the focus is conceptual or procedural, code as conceptual.

Procedural questions often but not always are "how to" questions.
How do you read -"27 ÷ 3 = 9?"
*How do you read "**π**?"*
Tell me in detail how you would use this ruler to measure this line segment.
What step do we do next to rename 13/4 as a mixed number?
Did you treat this proportion as if it were a fraction?

Procedural response: Responses state a rule or a procedure as, *"28 rounded to the nearest ten is thirty, because the rule says if the ones digit is 5 or more you round up."* Or, *"To multiply by ten, stick a zero at the end of the whole number."* Or, *"To add two fractions with unlike denominators you first have to rename each using a common denominator."*

Procedural task: Such a task only requires carrying out a procedure or reading a symbol. Completing six division problems, adding three fractions, completing a mad minute on basic facts. Also, drawing chips out of a bag in a probability lesson or spinning a spinner is a procedural activity. For most children after January of Grade 1, counting objects is a procedure. Doing long division examples with little or no reference to meaning of the steps is procedural. Measuring the sides of an object to find the perimeter, measuring an angle using a protractor.

\