

Does Subject Matter *Matter*?

by

Marilyn Chambliss and Anna Graeber

with Kathleen Clark

University of Maryland

DRAFT: Please do not quote or cite without the authors' permission.

Paper prepared for the annual meeting of the American Educational Research Association, Chicago IL, April 21-25, 2003. Direct all correspondence to Marilyn J. Chambliss at mc211@umail.umd.edu. The work reported herein was supported by the Interdisciplinary Educational Research Initiative (IERI), 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

A type of dualism exists for the study of high-quality teaching across subject matters. Historically, teaching has often been characterized according to either generic principles that ignore domain differences or domain-specific principles that ignore commonalities. This paper reviews recent research and scholarship in high quality teaching in reading and mathematics and reports preliminary results of an observational study of fourth and fifth grade teachers to explore where and how subject matter *matters*. The review highlights important differences in how scholars study effective teaching in the two subject matters. The data analysis reveals aspects of high-quality teaching that, even when similar at a generic level, play out differently in mathematics and reading. What appears to be similar on the surface may actually have quite a different epistemology, leading to different types of subject matter understandings for teachers and different types of learning for students.

Does Subject Matter *Matter*?

A picture of classroom life has emerged in this study that differs from other portrayals. Our investigation of fifth-grade math and social studies classes highlights the internal variety of the school day, not its sameness. We have found teachers do not teach in only one way; they alter their instructional approach depending on what they are teaching. Students do not learn in only one way; they are expected to approach mastery of different school subjects through different activities and tasks. When examining classroom activity, the subject matters for teachers and pupils (Stodolsky, 1988, p. 103).

The purpose of this paper is to explore how 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 adults teach both reading and mathematics, and, of course, the same children learn them. Indeed, Stodolsky (1988) discovered that instructional activities with high cognitive goals increased student engagement equally regardless of subject matter, particularly in cooperative groups where students had control over their own pacing. It is plausible that high quality teaching looks the same across reading and math.

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, analogous to Stodolsky’s (1988) findings in mathematics and social studies. To be sure, in teacher preparation programs across the country, different faculty teach elementary reading and math methods, and often a third cohort of faculty teach effective general teaching

practices divorced from specific subject matters. 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. Stodolsky (1988) recommended that the generalizability of the patterns in her study be examined for other teachers, schools, and subject matters. This paper builds on her work to review recent research and scholarship in high quality teaching in reading and mathematics and to report preliminary results of a study with a larger number of teachers in another part of the country chosen because their students' achievement in reading and mathematics is higher than would be expected given student SES.

For almost three years, we have had a compelling reason for identifying teaching practices that are common to reading and mathematics while remaining true to the many differences between the two subject matters. As Senior Researchers on the High Quality Teaching Study, we, one of us a reading educator and the other a math educator, have worked together to develop observation protocols and teacher logs that would measure effective practices of fourth and fifth grade teachers in moderate to high poverty schools. We have reviewed the literature in our respective fields, developed prototype measures, tried the measures out in classrooms, and revised accordingly.

Perhaps surprisingly, we discovered that many important aspects of teaching in our two subjects could be captured quite well by seven major categories: Teacher Activity, Student Activity, Classroom Organization, Content, Context, Classroom Behavior, and Technology/Materials. Of course, specific items for Content differed dramatically. Nonetheless, items to characterize what teachers and students did, how

instruction was organized, and so on, across reading and mathematics were often identical (e.g., Teacher requests an alternative answer; Teacher watches/reads student work). These many similarities suggest that the scholarship in reading and mathematics instruction may lead to similar conclusions; that high quality teaching in reading and mathematics may actually “look” quite similar. However, these similarities mask important differences in the research base of the two domains. As we began to craft this paper, we realized anew that we would lose much if we glossed over the epistemic differences between the study of instruction in reading and mathematics. Consequently, we decided to single author the major sections in this paper that review scholarship in each of the domains.

Our paper begins with a discussion of high quality teaching for reading as it has been studied by other scholars followed by a comparable discussion for mathematics. In a third section, we describe preliminary results from the High Quality Teaching Study in which we have designed our work to be sensitive to both the features of high quality teaching that reading and mathematics instruction have in common as well as the features that distinguish the two. All three of these sections are guided by the seven major categories that under gird the instrumentation in the High Quality Teaching Study. The final section returns to a discussion of the guiding question, “Does subject matter *matter?*”.

High Quality Teaching in Reading

High quality teaching in reading can be an elusive construct. Is high quality teaching in reading the instruction that promotes the literacy standards developed jointly by the International Reading Association (IRA) and the National Council of Teachers of

English (NCTE)? Is it instruction that matches the goals and teaching prescriptions designed by individual states and school districts? Is high quality teaching the set of practices used by teachers whose students score well on standardized tests? Is it instruction that leads to higher reading performance than other instruction offered to a control group? Is it instruction provided by teachers who have been recommended by their colleagues as outstanding teachers of reading? In a well-designed world, each of these approaches to identifying high quality teaching would reveal the same set of teaching practices. And, as this section will demonstrate below, the separate sets of teaching practices do overlap, sometimes substantially, at least in part reflecting the understandings that have been laid out in *Standards* (IRA & NCTE, 1996). However, to the extent that different understandings of teaching and learning outcomes underlie the work, the characterizations of high quality teaching in reading may also differ meaningfully. Because *Standards* was designed to reflect a shared understanding between researchers and teachers, this section begins by characterizing high quality teaching in reading based on these standards. See Table 1 for an overview.

Standards for the English Language Arts. *Standards* (IRA & NCTE, 1996) clearly characterized the knowledge and strategies that all students should develop in the English language arts. And, in general, *Standards* characterized the curricular and instructional features that would foster this development. See Table 1, Column 2, for an overview. However, an important viewpoint evident throughout the document was that there is no one curriculum or approach to teaching that will be equally effective with all students in all places. Students develop language in different ways and at different rates. Therefore, adaptability and creativity will be the hallmark of high quality teaching as teachers

respond to the learning needs that arise. Teachers are responsible for recognizing and valuing all children's rich and varied potentials for learning and for providing appropriate educational opportunities to nurture them. The needs of the learner are to be at the core of all instructional decisions.

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 education must therefore be to enable students to use language to address their own needs as well as the needs of their own families, communities, and society as a whole. This learning best takes place through social interaction within a particular discourse community. Students need frequent opportunities to read, write, talk, and listen, actively engaged in using language to communicate. They need to become acquainted with the wide variety of genres and modes that literate people use to communicate and the array of processes or strategies that will allow them to read and write successfully. In line with these general premises, *Standards* suggested specific types of instruction that would characterize high quality teaching.

High quality teaching can be recognized by the content that it covers. First, it introduces students to a wide range of print and nonprint text that represents fiction and nonfiction genres, classic and contemporary works, different difficulty levels ("easy" for fluency; "instructional" to develop skills), and diverse cultures. Second, it introduces students to and gives 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. Finally, it

includes 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).

High quality teaching can be recognized by particular instructional approaches. Students actively read, write, speak, listen, view, and visually represent texts for a purpose that is clear to them. The four purposes outlined in *Standards* are (a) obtaining and communicating information, (b) literary response and expression, (c) learning and reflection, and (d) problem solving and application. Within each of these purposes, students have frequent opportunities to talk and write as learners and thinkers. High quality teaching encourages and rewards 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 established conventions and ways of expression. These activities occur in social settings, (e.g., dyads and small groups) that encourage the active participation of all students.

It is important, however, to end this discussion of *Standards* with a reminder, which appeared throughout the document. The most important characteristic of high quality teaching is that it respects each student’s home language, prior knowledge, and cultural experience, and provides appropriate and rich instructional support accordingly. Consequently, high quality teaching in one classroom in one school might differ from high quality teaching in a classroom in a different school depending on the knowledge and language skills of the students in the class. It likewise might differ from one year to the next for the same teacher working with different students. The central focus in *Standards* was on the knowledge, skills, and processes that all students were to reach, and to a much lesser extent on the curriculum and instructional practices that would lead

to these goals. Specific curriculum and instructional decisions were to be left in the hands of local school systems and individual teachers.

Student success on standardized tests. High quality teaching is intended to lead to valued student outcomes. States and districts choose or develop mandated assessments to measure whether teachers' students are meeting the intended goals. Indeed, the results from large scale assessments have been used to identify the relationships among certain instructional practices and student outcomes. Sophisticated statistical analyses have been developed to isolate the effects of teaching from other effects such as a student's school, previous reading instruction, SES, and family background. Admittedly, it is beyond the scope of this paper to discuss the merits (and demerits) of each of these statistical approaches. Nonetheless, this subsection reports a recent reanalysis of the *Prospects* study that estimated both the overall magnitude of teacher effects in Title I schools on growth in student achievement, as well as the effects of particular instructional practices (Rowan, Correnti, & Miller, 2002). See Table 1, Column 3, for an overview. The reanalysis used two cohorts: those going from first to third grade over the course of the study and those going from third to sixth grade. This reanalysis offered a somewhat different perspective on high quality teaching from *Standards*.

First, the reanalysis revealed that differences among teachers were related meaningfully to differences in rates of student academic growth in reading. After controlling for student background variables, the classrooms to which students were assigned in a given year accounted for roughly 60-61% of the reliable variance in students' rates of academic growth in reading achievement (depending on the cohort), with d-type effect sizes ranging from .77-.78 (Rowan et al., 2002). Second, the more time

that students spent in whole group settings ($d=.09$) and the less time that they spent in individualized settings ($d=-.09$) were related to the amount of student academic growth in reading. Rowan et al. (2002) concluded that the most important difference between these two settings was the direct involvement of the teacher and noted that active teaching had a positive effect on students' growth in achievement comparable to results in earlier research. Interestingly, however, the particular type of active teaching did not seem to affect student outcomes differentially. The amount of time that teachers either presented or explained material, monitored student performance, lead discussion groups, or provided feedback on student performance did not have statistically significant effects on students' growth in reading (Rowan et al., 2002). Finally, content coverage also had an effect, although the analysis was only conducted for the lower grades. The effect of a teacher's emphasis on word analysis skills was $d=.10$, on comprehension was $d=.17$, and on writing was $d=.18$, suggesting the positive effects of a balanced approach to reading instruction (Rowan et al., 2002).

In comparison to the rich conceptualization outlined in *Standards*, this reanalysis provided a rather impoverished characterization of high quality teaching, as a comparison of Columns 2 and 3 display graphically in Table 1. To be sure, the outcomes suggested that differences in instruction can make a difference in students' reading achievement growth, and that this difference can be affected by how the teacher teaches (whole class versus individualized), and what the teacher covers (decoding, comprehension, and writing). But the focus on both high cognitive level and child-centeredness found in *Standards*, as well as in Stodolsky's (1988) work, was not evident in this reanalysis. Indeed, Stodolsky found that both teacher involvement and individualized work were

inversely related to cognitive challenge; that cognitively challenging tasks were most likely to be set in cooperative learning activities that were student led. While the exploration of large data sets can be revealing, meticulous observations of reading classes, similar to those conducted by Stodolsky in social studies and math, are necessary in order to determine whether the rich instructional experiences specified in *Standards* are present in classrooms and relate positively with student outcomes.

Comparisons of “treatment” and “control” groups. A third approach to identifying high quality teaching can be to try out a set of teaching practices with children in a “treatment” (or “innovation”) group and compare their subsequent performance with children in a “control” (or “comparison”) group. If the treatment group performs more competently than the control group and if the difference is statistically significant, then it can be concluded that the targeted teaching practices are features of high quality teaching. This section reports the results from a meta-analysis of experimental and quasi-experimental studies in reading that summarizes what can be learned from the studies that have been conducted.

In response to a Congressional charge in 1997, the Director of the National Institute of Child Health and Human Development in consultation with the Secretary of Education convened a national panel to assess the status of research-based knowledge about reading instruction (National Reading Panel, 2000). The panel subsequently conducted a meta-analysis of published experimental and quasi-experimental studies in five areas: alphabetics, fluency, comprehension, teacher education and reading instruction, and computer technology and reading instruction. See Table 1, Column 4, for an overview. The resulting report has been controversial (e.g., see Alvermann, 2002;

Pressley, 2002), and indeed includes a minority report. Nonetheless, it summarizes the features of high quality teaching in these five areas that have been identified through experimental work in which an instructional treatment is compared with a control.

Three of the five topics covered processes that characterize competent reading: alphabets, fluency, and comprehension, further subdivided into vocabulary instruction and text comprehension instruction. A fourth topic focused on the power of using technology to teach these reading processes. Overall, the results suggested that these processes can be taught and that certain instructional approaches are more efficacious than others, but due to variability in both constructs and measures for comprehension and the small number of studies for technology, results of the analyses were more statistically “rigorous” for alphabets and fluency than for these other topics.

Phonemic awareness instruction ($d=.53$), and phonics instruction ($d=.41$) had a positive effect on children’s decoding and to a lesser extent on their comprehension for all grade levels from pre-K through sixth grade (particularly for phonemic awareness), all SES levels, and all reading levels including at risk, disabled, and normal progress. The meta-analysis was complex. The following specifics are only suggestive of the range of moderator variables. For phonemic awareness, the effects were greatest when the instruction lasted from 10 to 18 hours ($d=.86$), dropping substantially from 20 to 75 hours ($d=.31$). The effect was also greatest for small group settings ($d=.81$) for phonemic awareness and tutoring ($d=.57$) for phonics, for children below second grade ($d=.55$) for phonics, and for children categorized as low SES ($d=.66$) for phonics, but mid and high for phonemic awareness ($d=.84$). Finally, guided repeated oral reading instruction had a positive effect on children’s reading ($d=.48$), both fluency and comprehension.

Because of the variability of both constructs and measures, the Panel was unable to compute effect sizes for different types of comprehension instruction, and because of the small number of studies, a meta-analysis was not possible for studies on the uses of technology. Nonetheless, in reviewing the literature, panel members concluded the following. Overall, a variety of methods is probably best for teaching vocabulary, but explicit instruction is more effective for low ability readers, and indirect instruction is more effective for high ability readers. Comprehension monitoring, cooperative learning, graphic organizers, question answering, question generation, story structure, and multiple strategy instruction have all been shown to affect text comprehension positively. Based on meta-analyses of reciprocal teaching cited but not conducted by the Panel, teachers are most effective who demonstrate, model, explain, and implement interaction. And the small number of studies using technology suggest the value of adding speech enhanced text, hypertexts with links to word definitions and other background information, and word processing.

As a comparison between Columns 2 and 4 in Table 1 suggests, the experimental work adds some research-supported specifics to the characterization of high quality teaching in *Standards*. For example, to create “instruction to build on students’ language proficiencies and experiences, teachers can ask questions, demonstrate, model, explain, and teach either explicitly or implicitly depending on students’ reading abilities (Teacher Activity on Table 1). Some of the “Strategies and processes” that teachers can teach are phonemic awareness, phonics, fluency, vocabulary, comprehension monitoring, and story structure (Content on Table 1). Teachers can promote dialogue among students by placing them into cooperative learning groups (Organization on Table 1). However, this

meta-analysis has been criticized for not exploring how practical many of these instructional approaches would be for real classrooms (Yatvin, 2000), and indeed the Report acknowledged that classroom implementation has been under researched. Alvermann (2002) has suggested that the Panel's conclusions be interpreted with caution because the Report did not include research on either second language reading or reading in the content areas and excluded all studies conducted using qualitative methods, thus limiting what can be known about the influence of different contexts. She suggested that six of the seven comprehension strategies are "...ones that teachers would use if they believe the reading process typically consists of students working individually to extract information from written texts" (p. 194).

Classroom observations of high quality teaching. A fourth approach to characterizing high quality teaching is to collect observations of teaching practices from teachers whose students score higher on standardized tests than would be predicted based on other indicators or who have a reputation among their colleagues as providing quality teaching. Observational studies are poised to characterize high quality teaching on a regular basis as it exists in real classrooms.

Primarily under the auspices of the Center for the Improvement of Early Reading Achievement (CIERA), a corpus of studies has focused on schools that serve students of poverty and have been particularly successful at teaching these students to read (Taylor & Pearson, 2002). Note that Rowan et al. (2002) found school effects ($d=.55$) as well as teacher effects in the prospects data. While scholars have developed a number of interesting case studies of successful schools (e.g., Smith, Johnson, & Jones, 2002), this paper focuses on two studies of teachers, one of accomplished third grade teachers within

successful schools (Taylor, Pearson, Clark, & Walpole, 2002) and the other of fourth-grade teachers who had been recommended by district supervisors and administrators to be particularly effective at teaching large numbers of students to be readers and writers (Morrow, Wamsley, Duhammel, & Fittipaldi, 2002). Out of the many studies on which this paper could have focused, these two examined classroom instruction closely, used different methodologies, and considered grade levels that were close to those studied by Stodolsky (1988) and currently being studied in the High Quality Teaching study.

Taylor and colleagues (2002) chose teachers in two phases. First, they selected schools based on reputation for having implemented reading reform and for promoting greater than expected primary-grade reading achievement. Principals from each of these schools nominated teachers in grades 1-3. Four students from each of the teachers' classes were chosen for testing: two average performing and two low performing readers as identified by the teachers. In the fall and again in the spring, students were administered the QRI-II, an informal reading inventory.

However, reputation resulted in quite a bit of variability on the researchers' measures among both schools and teachers. Consequently, Taylor et al. (2002) rated schools and teachers on scales from 1-3 based on additional posthoc analyses. Schools were rated according to student gain scores on both researcher administered measures and district administered achievement tests. The teacher accomplishment rating was determined by whether the teachers demonstrated instructional practices during time-sampling observations that have emerged from other research. Both school and teacher ratings were used as between-subject variables in the analyses with home communication, time on task, time spent in small-group/whole-group/independent

activity, time in independent reading, preferred teaching style of coaching/telling/recitation, time spent writing in response to reading, time spent reading aloud, type of word recognition instruction for grades 1 and 2, and approaches to comprehension instruction as dependent variables. See Table 1, Column 5, for an overview.

Accomplished teachers (with ratings of “3”) had the highest amounts of time when their students were on task and in small group instruction (with an adult), and were more likely to have a preferred interaction style of coaching than less accomplished teachers. In contrast, their students spent the least amount of time in whole group instruction or independent activity, and these teachers were least likely to have a preferred interaction style of telling. Survey questions revealed that teachers in the most effective schools formed groups based on reading ability, but monitored their groupings frequently based on results of an in-school based assessment system, and flexibly regrouped according to the outcomes. In comprehension instruction, more of the most accomplished teachers frequently asked text-based and higher order questions and had the children write about what they read, although comprehension instruction occurred less frequently overall for all three levels of teachers than word identification instruction. Furthermore, even accomplished teachers were more frequently observed asking text-based questions or having children write about what they read than asking higher order questions. It is important to note that both communication with parents and silent reading distinguished higher from lower performing schools, but did not distinguish among classrooms in schools. It seems plausible that these two activities, while important

features of high quality teaching, may be more likely to occur (or not) in every classroom in a school and not differentiate among a school's teachers.

While these outcomes are more informative than the results from the reanalysis of the prospects data and report on the daily practices of classroom teachers in contrast to the Panel's meta-analysis, they still portray high quality teaching as being less rich than the characterization in *Standards*. Particularly lacking are a consideration of the variety of genres and literacy activities that children would need to experience in order to participate as full-fledged members of a literate discourse community as well as evidence of the high levels of cognitive engagement present in the effective social studies and math classrooms in Stodolsky's (1988) work.

A different methodology with a different group of teachers at a higher grade level focused more closely on a classroom as a discourse community. See Table 1, Column 6, for an overview. Morrow and colleagues (2002) observed four fourth-grade teachers across three districts in one state. Teachers were observed repeatedly and interviewed. Four broad categories emerged from the observations: types of reading and writing, content area connections, classroom environments, and classroom management. Results from the four teachers were synthesized into a single case study and presented as if the four teachers were one.

The case study reported that students regularly engaged in a number of different types of reading and writing activities in a number of settings. Students worked alone on sustained silent reading, writing in literature response logs and journals, and answering questions about their reading, and writing that demonstrated a particular skill. They worked in groups to learn specific reading skills, to discuss books that they had chosen,

and to write synopses of books that they had read. The teacher formed two types of reading groups: guided reading groups that were organized according to reading level in which the teacher taught the students needed skills, and literature circles based on book choice and led by the students themselves. In whole class instruction, the teacher read aloud to the students, modeling how to predict, clarify story elements, use context cues to understand vocabulary, make connections between a text and their own knowledge and experiences, and read with expression. The teacher also worked with the whole class on spelling, grammar, and punctuation.

These fourth graders read and wrote a variety of genres including several types of poetry, plays, advertisements, and expository texts across the content areas. The classroom presented a rich literacy environment with bookcases and a book cart filled with hundreds of books at all levels and genres, including children's literature, science topics, and biographies for social studies. The walls were covered with posters that defined literary terms, presented publishing guidelines, exhibited students' written work, and presented words for word study. Easels and bulletin boards displayed various topics that invited children to provide written answers and suggestions. Students' desks were grouped in clusters to facilitate group work. The teacher orchestrated the literacy activities within this rich environment by encouraging student cooperation, engagement, and independence.

This final case study portrayed teaching that most closely matches the depiction in *Standards* (a variety of genres, a variety of literacy activities, attention to children's needs). It also had elements of the *Prospects* reanalysis (some whole class instruction and a variety of literacy activities) and Stodolsky's early work (cooperative learning

groups—the literature circles). However, this rich portrayal came with a price. The synthesis did not allow for a comparison among teachers or an appraisal of the effectiveness of each of these features in a literacy classroom.

Implications for the High Quality Teaching Study. The patterns in Table 1 suggest both the power and the usefulness of the categories that are under girding the classroom observation and teacher log instruments in the High Quality Teaching Study. First, these categories usefully organize the characterizations of high quality teaching that can be derived from scholarly works with different epistemological bases and, consequently, different methodologies: standards developed by professional organizations, a representative large-scale study, a meta-analysis of experimental and quasi-experimental studies, an observational study of accomplished teachers, and a case study. Second, while *Standards* provides characterizations for each of the categories, none of the other works addresses all of them. Note the spotty treatment of Technology, Materials, and Context, for example. Third, across the five characterizations, there are both overlap and differences. For instance, in all five characterizations, how the teacher organized the class is an important instructional feature, but the effectiveness of the different types of organization differed across characterization. Rowan et al. (2002) reported that the amount of whole class instruction was positively related to student gains while the amount of time in individualized instruction was negatively related, while Taylor and Pearson (2002) reported that the most accomplished teachers organized higher amounts of small group instruction and lower amounts of **both** whole group and individualized instruction than less accomplished teachers, a finding also suggested by Stodolsky (1988). And Morrow et al. (2002) described effective teachers as providing all

three organizations, depending on the type of student activity and lesson content, an outcome that is most closely aligned with the depiction in *Standards*, and that is suggested by the National Reading Panel Report (2000) as well.

As the introduction to this section stated, high quality teaching in reading can be an elusive construct. Research results can conflict. At least at the level of major categories, the High Quality Teaching Study would seem to be poised to untangle some of the mixed results and fill in missing pieces of the research landscape. In a final section of this paper, we provide an overview of the study and present some early findings that nonetheless are suggestive. But first, we focus on high quality teaching in mathematics.

High Quality Teaching in Mathematics

Just as in reading, the concept of high quality teaching in mathematics is elusive. While it might be tempting to say we recognize the results of high quality teaching, this too is much debated. The media's attention to the "math wars" often brings attention to what mathematics should be taught and how it should be taught. Media coverage of the mathematics reform arguments has less often explored the values that drive these conflicting views. Responses to what is considered high quality teaching cannot be divorced from the goals set for teaching mathematics. As educational researchers, it is both humbling and helpful to remember Hiebert's (1999) list of three things research cannot do. It cannot establish "the best" methods; it cannot imagine new ideas; and while it can influence our values in setting standards, it is not the only basis for establishing standards (pp. 4-9). While every faction involved in the disputes over mathematics will agree that students should be problem solvers who can use mathematics with understanding, there are different visions of the order of events in learning, the amount of

efficiency with written computational algorithms that is desired, which algorithms are worth learning to the point of automaticity, as well as the roles that inventing, practicing, and reasoning should play in learning mathematics at the elementary level. Thus, in examining what we “know” about the quality of teaching in mathematics or what we consider to be high quality teaching in mathematics, the backdrop of values and goals must always be taken into account.

Toward the end of her book, Stodolsky reminded her readers of the “existential fallacy, the assumption that knowledge of *what is*, tells us *what ought to be*” (1988, p.129-130). Thus it is important to look at both what is and what mathematics instruction ought to be. There are recommendations for how mathematics should be taught that emerge from professional teaching standards, such as the National Council of Teachers of Mathematics (NCTM) 1989, 1991, and 2000 documents; research results from the implementation of standards-based programs; research results from earlier “process/product” studies; insights derived from international studies of teaching, such as the recent TIMMS 1999 Video Study; and, of course, the opinions brought by practicing mathematicians and everyone who has been a student.

NCTM Standards. The 1989 document, *Curriculum and Evaluation Standards for School Mathematics* (National Council of Teachers of Mathematics), was an attempt on the part of the mathematics education community to present a vision of what it means to be mathematically literate at a time when technology carries out many mathematical procedures that are being applied in a growing number of fields. The standards also were to serve “to guide the revision of the school mathematics curriculum and its associated evaluation” (NCTM, 1989, p.1). The 1989 document set out content goals for students in

three grade levels, K-4, 5-8, and 9-12. Four process standards (problem solving, reasoning, connecting, and communicating) and nine or ten content standards (e.g., number sense, whole number computation, geometry, measurement, fractions and decimals, algebra, etc.) were described for each of these grade bands. Relatively few pages were devoted to describing the type of teaching that would lead to the proposed goals. However, the document did set out goals for mathematics, both in terms of societal goals and goals for students, as well as project a cognitive/constructivist philosophy of learning.

The societal goals (mathematically literate workers, lifelong learning, opportunity for all, and an informed electorate) suggested a broadening of the content goals, more attention to issues of equity, considered use of technology for computation, an emphasis on problem solving and reasoning, as well as increased attention to communication of mathematical ideas and working with others on mathematics (NCTM, 1989, pp. 4-5). The overarching goal set for students was to achieve mathematical power (that is, to learn to value mathematics, become confident in use of mathematics, become problem solvers, and communicate and reason mathematically). This goal reflected a view of mathematics as “more than a collection of concepts and skills to be mastered; it includes methods of investigation and reasoning, means of communication, and notions of context “(p. 5).

In discussing the nature of teaching and learning, the document argued that although practice of facts and algorithms traditionally came before problem solving, change was needed. Problem situations should be first, providing motivation and context for learning skills. The notion of “learners passively absorbing information, storing it in easily retrievable fragments as a result of practice and reinforcement,” was to be replaced

by the view that “learners approach a new task with prior knowledge, assimilate new information, and construct their own meanings” (p. 10). The document argued that this view of learning meant that “instruction should vary and include opportunities for project work, group and individual assignments, discussion between teacher and students and among students, practice on mathematical methods, and exposition by the teacher” (p. 10).

More specifics on the view of teaching appeared in the 1991 NCTM document, *Professional Standards for Teaching Mathematics*. The purpose of this document was to set forth an “image of mathematics teaching” that would support reaching the goal of mathematical power (NCTM, 1991, p. 1). The notion of mathematical power was expanded to include a disposition to use mathematics in problem solving. “Flexibility, perseverance, curiosity, and inventiveness” were also identified as attributes that impacted a student’s mathematical power. Other assumptions included “*what* students learn is fundamentally connected with *how* they learn it, all students can learn to think mathematically,” and “teaching is a complex practice not reducible to recipes or prescriptions” (pp. 21-22). In view of these goals emphasis was given to six aspects of teaching. One was the posing of worthwhile tasks (namely those that involve significant mathematics, are based on knowledge of students, engage intellect, promote reasoning and communication, etc.). Three of the standards highlighted the role of discourse in the classroom. This included the teacher’s role in leading discourse that demands, appreciates, and considers students’ reasoning; challenges students’ thinking; and promotes exchange of alternative ideas among students. Also included was the encouragement of use of technology, models, diagrams, and symbols as means of

enhancing classroom discourse. The fifth standard characterized the desired classroom environment, including valuing students' ideas, encouraging risk-taking, providing independent and group setting, etc. that would lead to mathematical power. The final teaching standard called for teachers to “engage in ongoing analysis of teaching and leaning” p. (63).

In 2000, NCTM published *Principles and Standards for School Mathematics*, the result of planned activity to revise and update the previous standards documents. The 2000 *Principles and Standards* were also designed to bring into a single document standards for content, teaching, and assessment. The document was organized into four grade bands, K-2, 3-5, 6-8, and 9-12. At all grade levels content was organized in five content areas (number and operations, algebra, geometry, measurement, data analysis and probability) and five process standards (problem solving, reasoning and proof, communication, connections and representation).

The notion was maintained that learning must be with understanding, building new knowledge from experiences and prior knowledge. This document also recalled the caution of the earlier documents (“teaching mathematics well is a complex endeavor, and there are no easy recipes” (p, 17)), recalled the six teaching standards from the 1991 document, and offered a “teaching principle” to guide instruction (“effective mathematics instruction requires understanding of what students know and need to learn and then challenging and supporting them to learn it well”(p. 17)). The text elaborated on some of the 1991 statements, noting for example that worthwhile tasks alone are not sufficient, but that how they are presented, discussed, and carried out without eliminating challenge are critical. The 2000 document was to be more carefully supported by research , and the

recently released document, *A Research Companion to Principles and Standards for School Mathematics* (Kilpatrick, Martin, & Schifter, 2003), was intended to anchor the 2002 Standards in research and theory.

A Research Companion included some introductory chapters on teaching and included a section of four chapters, each representing a different perspective on teaching and learning (cognitive science, situative, socio-cultural, and eclectic). A chief thesis in an opening chapter by Hiebert was that the results of the traditional, direct teaching model are well known: “students knowledge of foundational topics of elementary mathematics is limited” (2003, p.12) and “students’ current mathematical performance is that students are more proficient with the processes of calculating, labeling and defining than with other mathematical processes such as reasoning, communicating, conjecturing, and justifying” (p. 12). After presenting evidence to support the claim that the more ambitious goals set forth in *Principles and Standards for School Mathematics* were not unreasonable, Hiebert characterized similarities across alternative programs that have been successful in having students attain more ambitious goals. Included were: “building directly on students’ entry knowledge, providing opportunities for both invention and practice, focusing on the analysis of multiple methods for solving problems, and asking students to provide explanations” (p. 15). Siegler’s 2003 chapter provided references to research that supported teaching activities such as encouraging analytic thinking and the potential benefits of group interaction in learning. Greeno (2003) reported on research indicating how didactic teaching limits students’ views of mathematics, its usefulness, and themselves. He also gave research examples from programs that have students participate more constructively. These indicated the benefits of classroom discussion, the

exploration of alternative methods and representations, engaging students in mathematics connected to life outside the classroom, and meaningful use of technology. Such students were more apt to see the mathematics they learned in the classroom as relevant to life outside the classroom and to bring meaning to the mathematical tasks they faced. Forman cited research that indicated that the form of instruction not only influences children's learning process and outcomes, but also their motivation and beliefs about learning and the goals of learning (2003).

Research on reform-based curriculum. During the 1990's a number of federally and privately funded curricula were developed to reflect the NCTM Standards. In 1994 the U.S. Congress mandated a review of mathematics curriculum to determine those that were "promising" or "exemplary." In a 1999 report from OERI (U.S. Department of Education) ten curricula were declared exemplary or promising. The implementation of these programs has become a venue for the study of curriculum, classroom practice, and student achievement. Senk and Thompson (2003) have brought much of this information together in their book, *Standards-Based School Mathematics Curricula*. While much of the attention in the studies was on the curriculum materials, the curricula all shared an emphasis on conceptual understanding over procedural skill that was distinct from traditional programs' emphasis on procedural skill over conceptual understanding (Kilpatrick, p. 478, 2003). While there are methodological concerns about some of the studies, "the tendency of students in new curricula to perform at the same level as comparison students on standardized tests and perform at higher levels on specially designed tests" (p.483, Kilpatrick, 2003) seems reasonable. Since the standardized tests often do not cover the goals of these curricula, and classroom implementation was not

always carefully documented, it is not clear if all of the curricula's goals were achieved. However, an emphasis on conceptual understandings in classrooms that attempted to be "standards-like" seemed to have produced at least expected achievement on standardized measures.

Adding It Up. In 1998 the National Research Council established the Committee on Mathematics Learning at the request of the National Science Foundation. Among the committee's charges was "To provide research-based recommendations for teaching... for improving student learning and to identify areas where research is needed"(Kilpatrick, Swafford, & Findell, 2001, p. 3). In the resulting report, *Adding It Up*, mathematical proficiency was defined as consisting of five strands: conceptual understanding, procedural fluency, strategic competence, adaptive reasoning, and productive disposition. The discussion of teaching focused on the interactions among teachers and students around content in a given context (pp. 313-314). High quality instruction was taken to mean instruction that "focuses on important mathematical content, represented and developed with integrity, ... takes into account students' current knowledge and ways of thinking, ... is effective with a range of students and over time, and develops the knowledge, skills, abilities, and inclinations that we term mathematical proficiency" (p. 315).

Four classroom vignettes representing different goals for and styles of instruction were presented and reviewed in terms of their possibilities for learning, risks to learning, and demand on teacher knowledge and planning. The discussion that looked at factors such as content, opportunity to learn, task selection, discourse, and expectations made few definitive statements about practices that contribute to student achievement.

However, the use of challenging tasks in ways that support students without reducing the task to routines, increased emphasis on meaning in mathematics, reliance on logic of the discipline rather than teacher authority, the valuing of student ideas and alternative methods, the promise of problem settings for learning skills, and the use of cultural diversity to enrich learning experiences were discussed as promoting student achievement—at least in some settings. Here too we read that “effective teaching ... can take a variety of forms” (p. 315).

Process/product studies. The recommendations that are drawn from the NCTM Standards documents and the studies of curricula drawn from them stand in relatively sharp contrast to the direct instruction or “active teaching” models that evolved from the process-product studies of the 1970s. Although many of these studies involved primary grade students in reading or mathematics, the Good, Grouws and Ebmeier (1983) *Missouri Mathematics Effectiveness Project* is among the most widely known such studies, involved an experimental phase, and focused on mathematics learning at the fourth grade level. After identifying instructional behaviors associated with teachers whose students achieved at higher levels than those of other teachers, Good, Grouws, and Ebmeier conducted a treatment study in which the resulting model of effective teaching was shared with some teachers. The teachers’ implementation of that model was noted, and the researchers then compared the achievement of the treatment teachers’ students with that of a group of control teachers’ students. The treatment teachers’ students outperformed the control teachers’ students on both a standardized achievement test and a specially designed criterion-referenced test of material actually taught.

The “active teaching” model called for whole-class instruction, brisk pacing, structuring for high success rates, and homework assignments. Lessons were to include a daily review session, and a 20-minute development portion, during which the teacher was to explain or demonstrate new material and assess student comprehension. Whole class or independent activities were to be favored over small group work. The emphasis on quick pacing and high success was consistent with posing low-level questions with simple answers. Development of new ideas was seen as a teacher responsibility, and discussion or elaboration of student responses was not seen as conducive to student achievement. Further analysis of the data indicated that this approach was especially effective with low achieving students.

Critics of the process product studies of the direct instruction models noted that the studies were not based on theoretical models and that outcome measures used in the early studies reflected largely low-level factual information and procedural skill. Some (Becker & Gersten, date) argued that gains made by low achieving students in direct instruction programs were lost over time. Other research has suggested that higher-level problem solving, reasoning, and communication of mathematics demand a different type of instruction (see for example Doyle, 1983 and Peterson, 1979).

More recent recommendations concerning direct instruction have echoed those in the active teaching model. They have drawn on research such as that by Berliner, Stallings, Carnine, etc. For example, Stein, Silver and Carnine (1997), in their characterization of effective mathematics instruction, indicated that research suggests that a predominance of routine, low level tasks (p. 7), with fewer problems or demanding tasks (p. 36) is beneficial to achievement. Explicit strategy instruction in which every

step of a procedure is first prompted and prompts are then faded is cited also as beneficial to achievement. Stein, Silver and Carnine also advocated explicit statements of classroom objectives (p. 4) and a lesson with three phases: initial instruction, guided practice, and independent practice (p. 35).

International studies of teaching. The Second International Study of Mathematics (SIMS), the Third International Study of Mathematics and Science (TIMMS) as well as The TIMMS 1999 Video Study have been venues for researching the question of how classroom practice as reported or as observed contribute to student achievement. Methodologically the TIMMS 1999 video study (Hiebert et al., 2003), in which videotapes of classrooms were carefully analyzed, may be the most reliable source of such data. While not intended to examine best practice, one of the report's conclusions, "countries with the highest achievement teach in a variety of ways"(p.14), suggests that there is not one set of "highest quality" teaching practices, but rather that there may be different paths to high student achievement. Hiebert et al. (2003) noted that there are "complex interactions among features [of teaching] and the different kinds of learning that different configuration of features support" (p. 4).

Implications for instrument development. The standards-based and direct instruction suggestions for teaching imply some contrasting recommendations. While there are a myriad of models that could and are suggested as alternatives, the two shown in Table 2, represent stereotypical formulations that are at odds with one another in a number of ways. They are also largely premised on two different sets of values and goals for school mathematics. They also likely reflect two different views of mathematics. Putnam, Lampert, and Peterson (1990) suggest that mathematics includes both

“conceptual big ideas” and “conventions of mathematics” (p. 137). While the conceptual big ideas should be presented so as to foster student construction of them, conventions “must essentially be presented to students. . . . The line between these conventions of mathematics that must be acquired and the big ideas that must be constructed is by no means distinct”(p. 137).

Given the repeated admonitions that there is no unique pattern of teaching that assures outcomes in all situations for all goals, what did we consider in constructing classroom instruments? What patterns of teaching can we document? How did we incorporate research findings, theory and existing practice? Part of our strategy was to attempt to collect data that would capture facets of teaching important to either of these paradigms. Table 2 contains one view of the differences between two sets of desired teaching factors; it is not complete, but illustrative. Analogous to Table 1 in reading, the left-most column reflects the major categories in our High Quality Teaching observation instrument (but one of three major sources of data) that can capture classroom activity from either of the other two columns.

In coding episodes in classrooms, we can, for example, capture the extent to which a lesson includes demands for students’ reasoning or short answers. The level of tasks given to students is coded. The level of teacher questions and the nature of student discourse are also captured. In the arena of student behaviors, we collect data on the level of student responses, to whom students direct questions, and the level of tasks on which they work. Coding of classroom organization captures the extent to which whole class, small group, or independent work is used in the lesson. In coding content in mathematics, the specific topic is not coded, but the episode’s attention to conceptual knowledge,

procedural knowledge, or the linking of conceptual and procedural knowledge is captured. These are but a few of the examples that illustrate the way in which the instrument design facilitates gathering data on aspects of teaching relevant to different perspectives. Other observation responses were driven by what we saw actually occurring in classrooms, for example teachers posting information.

It is important to note that we did not include specific mathematics content in our observation protocol. There are numerous theories, studies and projects that detail what effective instruction might be in the case of teaching specific content (e.g., geometric concepts, understanding of basic facts, place value.) This pedagogical content knowledge often has to do with the order and sophistication of experiences and effective representations. While such knowledge is highly valued and important for consideration of teaching of the subject matter, the High Quality Teaching Study does not have as a goal the study of the use of such specific pedagogical content knowledge.

Implications for the High Quality Teaching Study. The extent to which teachers view some piece(s) of mathematics as convention may determine which of these two paradigms outlined in Table 2 are used. Teachers hearing both sets of recommendations would surely be pulled in different directions. However, as Croninger, Valli and Price (2003) have noted, the fact that currently there is emphasis on high-stakes tests that must be scored reliably and quickly, and that produce scores that are used to rate students and teachers, may well result in an emphasis on teaching for procedural efficacy. Other strands of mathematics proficiency outlined in the National Research Council's report, *Adding it Up* (that is adaptive reasoning, strategic competence, productive disposition, conceptual understanding) may be lost as teachers strive to fulfill district mandates

concerning pacing and prepare students for standardized tests that value procedural fluency and understanding basic concepts over adaptive reasoning. In the next section, we report on early findings from the High Quality Teaching Study.

Early Findings: High Quality Teaching Study

The High Quality Teaching Study is a five-year, longitudinal study of fourth and fifth grade teachers. The study investigates how teachers in moderate- to high-poverty schools scale up their practice to promote high levels of learning in reading and mathematics. The part of the study that we are presenting addresses two research questions:

What do teachers do to help fourth and fifth grade students achieve above predicted rates in reading and mathematics?

What are the attributes of high quality teaching as identified through collegial nominations, classroom observations, and achievement gains?

To answer these questions, we are conducting multiple classroom observations each year, collecting daily teachers' logs, collecting student records of annual achievement gains, conducting on-site interviews, observations, and school surveys, and conducting case studies of teaching practices. For this paper, we focus on early results from the classroom observations of teachers who were chosen for the study because they taught in schools with moderate to high poverty enrollments and higher than expected student performance on one or more indicators of achievement.

During spring 2002, the first or base-year of the study, we piloted observation protocols in 40 reading classes and 45 mathematics classes in 11 schools. The protocol used a three-minute time frame programmed onto a laptop computer. Every three minutes, the observer scanned the classroom and coded Teacher Activity, Student

Activity, Organization of Class, Lesson Content, Context (connections to background knowledge and interest), and Student Behavior, Technology, and Materials. To illustrate the detail under each category, Table 3 shows the coding subcategories under Teacher Activity and Student Activity for mathematics and reading. Items that only occur for one subject are bold-faced. This array of items allows us to characterize features of high quality teaching that are unique to a subject matter as well as those that the two subject matters share.

Results show both similarities and differences between mathematics and reading. Teacher Activity, Student Activity, Context, and Classroom Behavior had almost identical proportions of coding for the two subject matters. Table 4 shows the results for the top three major categories of teacher and student activity for mathematics and reading. Teachers were coded as establishing a context for instruction 17% and 15% of the time for reading and math respectively. Finally, students in both content areas tended to be on task for 82% of the codings.

Different patterns emerged for Content, which we expected, and both Classroom Organization and Technology, which initially surprised us. Perhaps the most interesting contrast was in Classroom Organization. Table 5 shows that over half of the codings for both reading and mathematics were “Whole class instruction/discussion”. However, the second most common coding for reading was “Independent,” followed by “Mixed small group and independent,” and “Small Group.” Math demonstrated a different pattern with “Small Group” the second most common coding, followed by “Independent” and “Mixed small group and independent”. For reading, the “Mixed small group and independent” category suggested that students in reading were taking turns working in a small group

with their teacher while the rest of the class sat at their seats working independently, perhaps allowing teachers the opportunity to adjust their instruction to match the literacy needs of the students in the small group, an approach that matches the recommendations in *Standards* (IRA & NCTE, 1996). However, if one includes the students who were working independently in the “Mixed” category, students in reading were spending far more time working independently than students in math, an interesting conundrum for a subject matter that values discourse and social interaction around text.

The contrast for the use of technology (22% of the codings for math; 2% for reading) may well be the result of differences in how the two subject matters both define and then use technology. For reading, the choices were confined to the use of computers, audiotape, or videotape. For mathematics, technology use was defined more broadly to include manipulatives and tools as well as various types of electronic technology.

As informative as these individual outcomes are, it may very well be the relationships among the individual characteristics that will turn out to be the most interesting. Table 6 presents the five strongest correlations between Teacher and Student Activity for mathematics and reading. Note that they range from moderate to strong (Guilford, 1956). The first conclusion that we draw from this analysis is that teachers “got what they asked for.” The first and fifth correlations on the table for math and reading clearly illustrate this relationship. Interestingly, the type of teacher activity in mathematics that was the most closely associated with students conjecturing, explaining/justifying, or stating an alternative method was the request to elaborate. The second conclusion is that the patterns in the two subject matters are not the same, an outcome that we find puzzling, but interesting.

It is important that we end this section with a reminder. These outcomes are based on pilot data from one of the several types of data that we are collecting. As such, they demonstrate the potential power of the observation protocol to replicate and build on current characterizations of high quality teaching. However, these pilot data are only a teaser for what is to come out of the multiple teacher observations and other types of data that we are collecting this year and will continue to collect for two more years. Furthermore, these data have not been related to student outcomes. We certainly can compare them to the work of others that we have reported, but at this point we do not know what the relationships among these data and student outcomes will be.

Does Subject Matter *Matter*?

All three previous sections of this paper suggest that the answer to our question at this point in time must be a qualified “yes.” Certainly, differences in the reading and mathematics sections of this paper suggest differences in how the Senior Researchers in reading and mathematics respectively are framing the issues in their two subject matters, and early data from the High Quality Teaching Study suggest some interesting differences. Other comparisons between mathematics and a second subject matter have found important differences as well. Stodolsky (1988), whose study of mathematics and social studies classrooms was quoted above contrasted the teaching of mathematics in which instruction varied little either within a teacher’s classroom or across teachers to the teaching of social studies wherein instructional variety prevailed (Stodolsky, 1988). She directly pointed to differences in teaching in various subject matters when she stated “individual teachers vary instructional arrangements as a function of the subject they teach” ... “teachers do not have one style of teaching” (p.131). 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 outcome may suggest that students too have different views of good teaching for different subject areas.

As we noted at the beginning of this paper, it is also possible that high quality teaching in reading and mathematics share similarities. Stodolsky (1988) suggested that the differences between mathematics and social studies may have resulted from the different levels of responsibility that teachers feel in teaching a "basic" subject matter (mathematics) over an "enrichment" subject matter (social studies). However, both reading and mathematics are basic subjects in elementary school. It is possible that high quality teaching in these two basic subject matters would share more similarity than do mathematics and social studies. More recently, the reanalysis of the data from the *Prospects* study by Rowan and colleagues (2002) contrasted mathematics and reading directly. On the one hand, they found only moderate correlations between the effects of teaching on student gain scores in reading and mathematics and therefore concluded, "a

given teacher varies in effectiveness when teaching different academic subjects” (p. 1533). On the other hand, Rowen and colleagues (2002) found some identical effects. For example, the same whole class instruction effect that they found in reading also was present in the mathematics data ($d=.12$).

Certainly, the reading and mathematics sections in this paper suggest some differences in the epistemology of the two subject areas. Our understanding of these differences is nascent at present. We offer some possible explanations without further comment. First, mathematics is not only a subject matter, but also a highly structured discipline. While it might be possible to understand reading as a discipline, that is not currently the case among scholars. Second, research in the two subject matters has different grain sizes that make it possible to compare and contrast across different types of scholarship in reading, but not as readily in mathematics. Scholarship in reading can be categorized according to large areas, i.e., alphabets, vocabulary, comprehension, that can be used to understand a variety of different types of scholarship from standards to meta-analyses, from large-scale studies to case studies. Except for the gross categorization of concepts, computation, and problem solving, these sorts of large categories do not obviously exist in mathematics. Research studies often focus on a very specific way to teach a very specific outcome, and it can be difficult to use the same categorization scheme across different types of scholarship (standards, large-scale studies, and case studies, for example). And, mathematics education research studies that look more broadly at elementary mathematics instruction involve various combinations of characteristics listed in Table 2, that are rarely if ever the same or attended to with the same proportion of attention given to particular characteristics (e.g., see Senk and

Thompson, 2003). Third, it may be the case that the goals of mathematics and means of getting to those goals are more divisive than the goals and means in reading by fourth and fifth grade. The divisiveness in reading that often occurs at lower grade levels may have been reconciled by upper elementary school.

The early results from the High Quality Teaching Study also suggest that subject matter *matters*. In addition to differences in content, the results highlighted different patterns in classroom organization and the use of technology. Even in cases where items share identical wording for the two subject matters, they may be measuring something different. Requesting that students elaborate on an answer in mathematics, for example, was highly correlated with high level responses from students, while this same relationship did not hold for reading. Requesting an elaboration in mathematics may require more depth of disciplinary understanding than it does in reading for both the teacher and the student providing the response. It is possible that even when high quality teaching in mathematics and reading looks quite similar, what appears to be similar actually has quite a different epistemology, leading to different types of subject matter understandings for teachers and different types of learning for students. In the future, a much larger data set will allow us to explore such differences more exactly.

References

- Alvermann, D. E. (2002). Effective literacy instruction for adolescents. *Journal of Literacy Research, 34*, 189-208.
- Becker, W.C., & Gersten, R. (1983). A follow-up of Follow Through: The later effects of the direct instruction model on children in fifth and sixth grades. *American Educational Research Journal, 19*, 75-92.
- 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 meeting of the American Educational Research Association, Chicago.
- Doyle, W. (1983). Academic work. *Review of Educational Research, 53*, 159-200.
- 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.
- Foreman, E. (2003). A sociocultural approach to mathematics reform: Listening, inscribing, and doing mathematics within communities of practice (pp. 333-352). In J. Kilpatrick, W.G. Martin, and D. Schifter (Eds.) *A research companion to Principles and Standards for School Mathematics* (pp. 304-332). Reston, VA: National Council of Teachers of Mathematics.
- Good, T.L., Grouws, D., & Ebmeier, H. (1983). *Active mathematics teaching*. New York: Longman.
- Greeno, J. (2003). Situative research relevant to standards for school mathematics. . In J. Kilpatrick, W.G. Martin, and D. Schifter (Eds.) *A research companion to*

- Principles and Standards for School Mathematics* (pp. 304-332). Reston, VA: National Council of Teachers of Mathematics.
- Guilford, J. P. (1956). Fundamental statistics in psychology and education. New York: McGraw Hill.
- Hiebert, J. (1999). Relationships between research and the NCTM Standards. *Journal for Research in Mathematics Education*, 30(1),3-19.
- Hiebert, J. (2003). What research says about the NCTM Standards. In J. Kilpatrick, W.G. Martin, and D. Schifter (Eds.) *A research companion to Principles and Standards for School Mathematics* (pp. 5-23). Reston, VA: National Council of Teachers of Mathematics.
- Hiebert, J., Gallimore, R., Garnier, H., Givvin, K., Hollignsworth,H., Jacobs, J., et al. (2003). *Teaching mathematics in seven countries: Results from the TIMSS 1999 video study*. Washington, DC: National Center for Educational Statistics.
- IRA & NCTE (1996). *Standards for the English language arts*. Published by the International Reading Association and the National Council of Teachers of English.
- Kilpatrick, J. (2003). *What works?* In S. Senk & D. Thompson (Eds.) *Standards-Based school mathematics curricula: What are they? What do students learn?* (pp. 471-488). Mahwah, NJ: Lawrence Erlbaum.
- Kilpatrick, J., Swafford, J., & Findell, B. (Eds.). (2001). *Adding it up: Helping children learn mathematics*. Washington, DC: National Academy Press.

- Kilpatrick, J., Martin, W.G., & Schifter, D. (Eds.) (2003). *A research companion to Principles and Standards for School Mathematics*, Reston, VA: National Council of Teachers of Mathematics.
- Morrow, L. M., Wamsley, G., Duhammel, K., & Fittipaldi, N. (2002). A case study of exemplary practice in fourth grade. In B. M. Taylor & P. D. Pearson (Eds.), *Teaching reading: Effective schools, accomplished teachers* (pp. 289-307). Mahwah, NJ: Erlbaum.
- 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.
- National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: National Council of Teachers of Mathematics.
- 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.
- Peterson, P.L. (1979). Direct instruction reconsidered. In, P.L. Peterson and H.J. Walberg (Eds.) *Research on teaching: Concepts, findings, and implications* (pp. 57-69). Berkely, CA: McCutchan.

- Pressley, M. (2002). Effective beginning reading instruction. *Journal of Literacy Research, 34*, 165-188.
- Putnam, R. Lampert, M., & Peterson, P. (1990). Alternative perspectives of knowing mathematics in elementary schools (pp. 57-150). In C. Cazden (Ed.) *Review of research in education: Vol. 16*. Washington, DC: American Educational Research Association.
- 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.
- Senk, S., & Thompson, D. (Eds.).(2003). *Standards=Based school mathematics curricula: What are they? What do students learn?* Mahwah, NJ: Lawrence Erlbaum.
- Siegler, R. (2003). Implications of cognitive science research for mathematics education (pp. 289- 303). In J. Kilpatrick, W.G. Martin, and D. Schifter (Eds.) *A research companion to Principles and Standards for School Mathematics* (pp. 5-23). Reston, VA: National Council of Teachers of Mathematics.
- Stodolsky, S. S. (1988). *The subject matters: Classroom activity in math and social studies*. Chicago/London: The University of Chicago 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.

U.S. Department of Education, Mathematics and Science Expert Panel. (1999).

Exemplary and promising mathematics programs. Washington, DC: Author.

Yatvin, J. (2000). Minority view. *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.