WCER Working Paper No. 2008-1
February 2008

What Do We Know About Instruction From Large-Scale National Surveys?

Eric M. Camburn
Consortium for Policy Research in Education
Wisconsin Center for Education Research
University of Wisconsin–Madison
ecamburn@education.wisc.edu

Seong Won Han
Consortium for Policy Research in Education
Wisconsin Center for Education Research
University of Wisconsin–Madison
seonghan@wisc.edu

WCER
Wisconsin Center for Education Research
School of Education • University of Wisconsin–Madison • http://www.wcer.wisc.edu/

WCER working papers are available on the Internet at http://www.wcer.wisc.edu/publications/workingPapers/index.php. Recommended citation:

> Camburn, E. M., & Han, S. W. (2008). *What do we know about instruction from large-scale national surveys?* (WCER Working Paper No. 2008-1). Madison: University of Wisconsin–Madison, Wisconsin Center for Education Research. Retrieved [e.g., March 1, 2008,] from http://www.wcer.wisc.edu/publications/workingPapers/papers.php

The research reported in this paper was supported by the Atlantic Philanthropies, USA; the William and Flora Hewlett Foundation; the U.S. Department of Education; the National Science Foundation; and the Wisconsin Center for Education Research, School of Education, University of Wisconsin–Madison. Any opinions, findings, or conclusions expressed in this paper are those of the authors and do not necessarily reflect the views of the funding agencies, WCER, or cooperating institutions.

What Do We Know About Instruction From Large-Scale National Surveys?

Eric M. Camburn and Seong Won Han[1]

Overview

Students' instructional experiences—that is, their experiences learning subject matter through classroom instruction—are a major determinant of how well they learn. Given the importance of classroom instruction, valid, generalizable evidence is needed by policymakers, researchers, and practitioners. But building a solid knowledge base on instruction has proven an elusive goal. More than a decade ago, Cuban (1993) argued that researchers, policymakers, practitioners, and reform-minded citizens know little "about which instructional practices have remained stable and which ones have changed over the last century" (p. 12). More recently, Ball and Rowan (2004) argued that our knowledge about what makes for good or effective instruction is quite limited.

Surveys of teachers and students are commonly used to measure instruction. Such surveys hold promise for shedding light on how instruction varies across settings and how different instructional practices affect achievement. Moreover, when administered to probability samples of students or teachers, surveys can yield generalizable knowledge on instruction. However, despite the widespread use of surveys to measure instruction in education research, to our knowledge there has not been an attempt to systematically document how evidence from surveys administered to probability samples has contributed to generalizable knowledge on instruction. This paper takes a step toward filling that gap by describing empirical evidence on instruction collected from large-scale surveys administered by the National Center for Education Statistics (NCES). Our goal in preparing this summary is to produce a portrait of the literature that highlights instructional activities that are positively associated with student achievement and identifies gaps in the literature that need further attention.

Background

Surveys are commonly used in education research to measure instruction and a range of other key constructs. In a recent bibliographic search, we identified 102 journal articles published in the 2006 calendar year that presented empirical evidence from teacher surveys.[2] The studies were conducted in a variety of countries (the U.S. and 25 others) and spanned all levels of the education system from kindergarten through postsecondary education. In these articles, surveys of teachers were most commonly used to measure psychological constructs such as beliefs, perceptions, and attitudes. However, 22 of the 102 articles focused on some aspect of

---

[1] We wish to thank Sarah Lubienski and John Smithson for helpful comments on an earlier draft of this paper. We also wish to thank Jimmy Sebastian and Chul Lee for their research assistance.

[2] To gain an understanding of how frequently teacher surveys were used in education research, we searched five major bibliographic databases (Academic Search Elite, ERIC, PsycARTICLES, PsycINFO, and SocINDEX) for the phrases *teacher survey* and *teacher questionnaire*. The search was limited to articles published in peer-reviewed journals between January 2006 and December 2006. This search yielded a total of 102 articles. We determined whether each article actually contained empirical evidence from teacher surveys by reading the abstracts.

instructional practice. Such frequent use of surveys suggests a potentially large body of generalizable evidence on instruction and psychological variables measured by the surveys.

In education research, much of the generalizable knowledge we have about instruction based on quantitative evidence comes from research summaries. Shadish, Cook, and Campbell (2002) pointed out that such summaries can either take the form of structured narrative summaries or involve systematic quantitative analysis using meta-analysis techniques. The basic logic behind these kinds of summaries is that one can have greater confidence that an empirical relationship (for example, the effect of a particular teaching strategy on student achievement) holds in general if that relationship is consistently observed across multiple study conditions, social contexts, and populations.

Since 1998, three notable research summaries have been published on instruction in the elementary grades in the U.S. In 1998, the National Research Council (NRC) published a narrative summary of research on reading instruction entitled *Preventing Reading Difficulties in Young Children* (Snow, Burns, & Griffin, 1998). In 2001, the NRC published *Adding It Up: Helping Children Learn Mathematics* (NRC, 2001), which includes significant narrative summaries of research on the kinds of instructional strategies that promote mathematical proficiency. In 2000, the National Institute of Child Health and Human Development (NICHD) published a report of the National Reading Panel (NRP) entitled *Teaching Children to Read: An Evidence-Based Assessment of the Scientific Research Literature on Reading and Its Implications for Reading Instruction* (NICHD, 2000). The NRP report contains a comprehensive summary of quantitative research on five areas of instruction: phonemic awareness, phonics, guided oral reading, vocabulary comprehension, and text comprehension. Of these three research summaries, the NRP report provides the strongest generalizable quantitative evidence on instruction since it is the only one to provide meta-analysis results, though the number of available studies permitted meta-analysis for only three of the five areas of instruction.

While research summaries support generalization through the cumulation of evidence across multiple studies, the results of studies utilizing probability samples are directly generalizable (Shadish, Cook, & Campbell, 2002). With formal probability sampling, study participants are selected with a known probability from a clearly defined population. Probability sampling is expected to produce comparable distributions between the sample and the population on all measured and unmeasured variables, within margins of sampling error. Given these kinds of comparable distributions, researchers are able to validly draw inferences from their sample data that generalize to the larger population (Shadish, Cook & Campbell, 2002). Perhaps the clearest example of this approach in education research is a series of large-scale surveys conducted by NCES during the past two decades. NCES has recognized the need for stronger evidence on instruction by measuring this important construct in teacher and student surveys. The NCES surveys, most of which employ nationally representative probability samples of students, allow for generalization to national student populations (typically to particular grade levels) and to a wide range of school settings. Each of the four NCES surveys examined for this paper utilized a multistage probability sampling procedure that first randomly selected schools and then randomly selected students within schools.[3] Given this sampling strategy, the studies

[3] The Third International Math and Science Study also included a middle stage between school and student selection in which classrooms within schools were randomly selected.

permit generalization only to various populations of U.S. students, not to the population of U.S. teachers.

Because the measures of instruction produced by these surveys are embedded within comprehensive, scientifically rigorous research designs, the surveys have the potential to yield valid and reliable evidence about how instruction is related to other important variables, most notably student achievement. Viewed as a whole, the NCES surveys have produced an impressive body of empirical evidence that has the potential to deepen our understanding of typical patterns of instruction in U.S. classrooms, factors affecting how teachers teach, and the impact of instruction on student outcomes.

This paper summarizes 98 studies that used data from the four large-scale NCES surveys to investigate classroom instruction. To date, this research has not been systematically analyzed as a unitary body of evidence. We believe that a comprehensive portrait of this body of generalizable evidence will provide a useful service to education researchers and policymakers by documenting areas of strength and potential gaps in the empirical literature on instruction. Our summary of this research is presented in two parts. We first describe general characteristics of this body of evidence using a database created for the study. We then summarize the results of a subset of 27 studies that used inferential statistics to examine instruction. We begin with a description of our research approach and conclude the paper with reflections on the implications of our findings for research on instruction.

Research Approach

The research summary presented here is based on analyses of a database of studies that used data from four major NCES surveys to examine instruction. In this section, we describe how the database was created and how it was used to conduct the research summary.

The database was created in three steps: (a) studies were selected for review, (b) each study was classified according to an a priori coding scheme, and (c) written summaries of the methodology and findings of each study were produced. The database thus includes the classification codes and written summaries. We discuss each step of this process in greater detail below.

All studies in the database used data from one or more of the following NCES surveys:

- The National Education Longitudinal Study of 1988 (NELS:88);
- The National Assessment of Educational Progress (NAEP);
- The Trends in International Mathematics and Science Study (TIMSS; formerly known as the Third International Mathematics and Science Study); and
- The Early Childhood Longitudinal Study (ECLS).

We chose these four surveys because (a) they used probability samples, (b) they focused to a significant extent on instruction, and (c) their data have been widely used by the education research community, resulting in a substantial body of research from which to draw.

We identified studies for inclusion in the database through document searches conducted in May and June 2005 using the Academic Search Elite and ERIC databases. We limited our search to studies published after 1987. The search engines were queried using a combination of study identifiers and keywords or phrases. Study identifiers included the full titles of the four NCES surveys (*Early Childhood Longitudinal Study*, *National Education Longitudinal Study of 1988*, *National Assessment of Educational Progres*s, *Trends in International Mathematics and Science Study*, and *Third International Mathematics and Science Study*) and the acronyms for these studies (*ECLS*, *NELS:88*, *NAEP*, and *TIMSS*). Keywords and phrases for which we searched included *instruction*, *teaching*, *instructional practice*, *classroom practice*, and *classroom activity*.

We retrieved additional documents through a manual search of the NCES Web site. The NCES Web site posts lists of publications for each of the four surveys we investigated. We read the titles and abstracts of all publications listed on the NELS:88, ECLS, NAEP, and TIMSS Web sites, downloading those that appeared to investigate instruction. We then read the downloaded publications to confirm whether they met our primary criterion of including an empirical investigation of instruction.

This initial search process identified a total of 565 studies—204 from TIMSS, 38 from ECLS, 298 from NAEP, and 25 from NELS:88. Of these 565 studies, 98 contained an empirical analysis of instruction based on data from one of the four NCES surveys and were analyzed for this research summary. We identified an additional four studies that fit all of our selection criteria but used data collected in countries other than the U.S. Since our goal for this research was to characterize generalizable evidence on instruction in the U.S., we did not include these four studies in descriptive analyses of the database. However, we reviewed them, and we mention them in the Narrative Summary of Research section below, noting that they used non-U.S. samples.

We next read each of the 98 studies in their entirety and coded them using a rubric designed to measure their scientific quality and their coverage of various aspects of the instructional system. We assessed *coverage* by coding which grade levels, subject areas, and dimensions of instruction the publication addressed. We documented *scientific quality* by coding the analytic approaches used and the venue in which the study was published. We discuss these study attributes in greater detail below.

Indicators of Coverage

Grade level and subject area. The quality of the research base on instruction depends in part on how well it covers the important dimensions along which instruction varies. Clearly, instruction varies considerably across grade levels and subject areas. The four large-scale NCES surveys examined in this paper cover a wide range of grade levels (K–12) and subject areas (English/language arts, mathematics, sciences, and social studies), though individual studies using the survey data typically focused only on one grade level and one subject area.

Dimensions of instruction. In our view, the quality of the research base on instruction also depends on how well it covers different aspects of instruction. We sought a generic framework for classifying studies with respect to the dimensions of instruction they addressed. Reigeluth and Moore's (1999) framework for comparing instructional strategies provided a useful scheme for this purpose. Their framework contains six dimensions that characterize an instructional system, theory, or approach:

1. *Type of learning.* This dimension of instruction characterizes the kinds of cognitive demands placed on students during instructional activities. Reigeluth and Moore (1999) identified four general types of learning that instruction might promote: (a) memorization, (b) application of specific skills, (c) understanding of relationships, and (d) application of generic skills. However, in coding studies for this dimension we did not limit ourselves to these four types.

2. *Focus of learning*. This dimension of instruction relates to the knowledge domains and topics on which instruction focuses. Reigeluth and Moore (1999) suggested that the focus of learning can be topic-oriented, problem-oriented, domain-specific, interdisciplinary, or some combination of these.

3. *Control of learning*. The third dimension of instruction focuses on the degree to which the pace of instruction is controlled by teachers, students, or both. Control can be exercised by providing choice over such things as the topics studied, the activities in which one engages, the pacing of activities, and the materials used.

4. *Grouping for learning*. This fourth dimension of instruction addresses the various ways in which students are grouped for instruction. For example, instruction can be individualized or provided to pairs of students, groups of students, or an entire class.

5. *Interactions for learning.* Reigeluth and Moore (1999) separated student interactions into two categories: (a) interactions with other people and (b) interactions with non-human resources. Interactions with other people include interactions between students and between students and teachers. Non-human resources with which students interact can include texts, tools (e.g., computers, manipulatives), and environmental resources (e.g., a zoo).

6. *Support for learning*. Reigeluth and Moore (1999) identified two general categories of support a student can receive through instruction: cognitive support and emotional support. Cognitive support consists of elements that serve to support students "in building their understandings of, and competence in, subject matter" (p. 64). Emotional support can consist of support for student attitudes, motivation, and feelings.

We added a seventh dimension to the Reigeluth and Moore (1999) framework: *time on learning*. The amount of time students and teachers spent on learning was a major focus of many of the studies we reviewed and struck us as conceptually distinct from Reigeluth and Moore's other dimensions.

Indicators of Scientific Quality

In addition to describing how well the studies we reviewed covered various aspects of instruction, we also examined two indicators of scientific quality—analytic approach and publication venue.

Analytic approach. Given our interest in assessing the state of generalizable knowledge on instruction from surveys, we found it useful to code whether studies used descriptive or inferential statistics. While descriptive statistics simply describe patterns in the data, inferential statistics allow researchers to formulate conclusions that reach beyond the data set being analyzed. In this sense, inferential statistics provide a stronger basis for generalization than descriptive statistics. Among the studies we reviewed, some used univariate descriptive statistics (e.g., frequency distributions, percentages, means), others used multivariate descriptive statistics (e.g., cross-tabulations), and still others used inferential statistics (e.g., *t* tests, correlations, multiple regression, and multilevel models).

Publication venue. Publication venue was our second indicator of scientific quality. For this paper, we examined articles published in refereed journals, books, published reports, and conference papers. We view the relative proportion of studies published in refereed journals as a gross indicator of the scientific quality of the empirical evidence base on instruction.

Characteristics of the NCES Studies Reviewed

In this section, we report on the characteristics of the studies in our database based on the document attribute codes just described. Of the 98 studies that used data from the four NCES surveys to empirically investigate instruction, 29 used TIMSS data; 48, NAEP; 14, NELS:88; and 7, ECLS (Figure 1). Given that we were looking at studies conducted over a 20-year period, the evidence on instruction yielded by the four NCES surveys comes from a wide range of years (Figure 2). Interestingly, only about one third of the studies used data from the past decade. Indeed, nearly two thirds (58%) used data collected before 1996.[4] This preponderance of older studies raised a question about the potential of the body of research to address current educational issues and challenges.

[4] Some studies used data from multiple years. To calculate the percentage of studies that used data collected before 1996, we counted the number of studies for which the *most recent* data set used was collected in 1995 or earlier. For example, if a study used data from 1988 and 1998, the study was counted as using data from 1998. Using this counting method, we found 57 studies that used data collected prior to 1996.

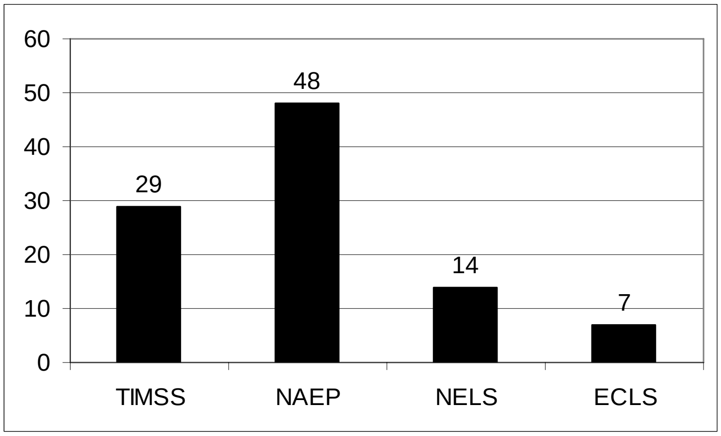


Figure 1. Number of studies reviewed, by NCES survey.

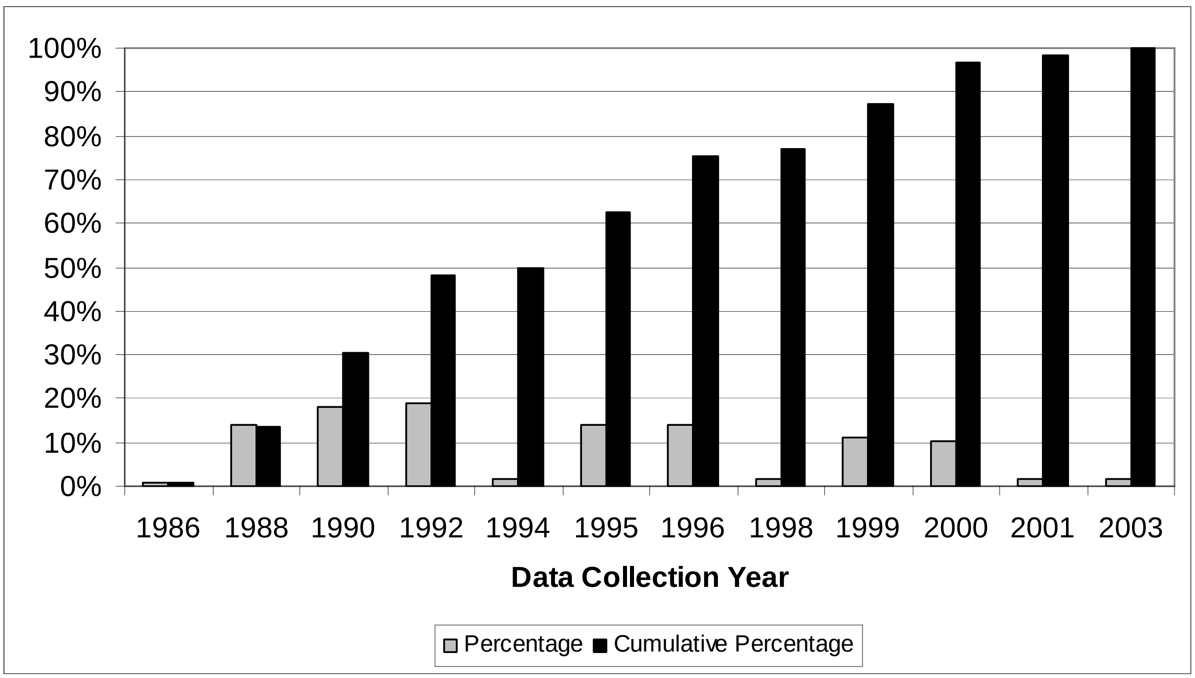


Figure 2. Percentage and cumulative percentage of studies, by data collection year.

Nearly half (48%) of all studies utilized data from annual teacher surveys *and* annual student surveys (Figure 3). The use of these two data sources took two distinct forms: in some cases, authors used data from the two sources to examine separate constructs, and in others, they used data from the two sources to examine two different perspectives on the same construct. Of the studies we reviewed, 38% relied solely on evidence from teacher surveys, whereas only 14% relied solely on student questionnaire data.

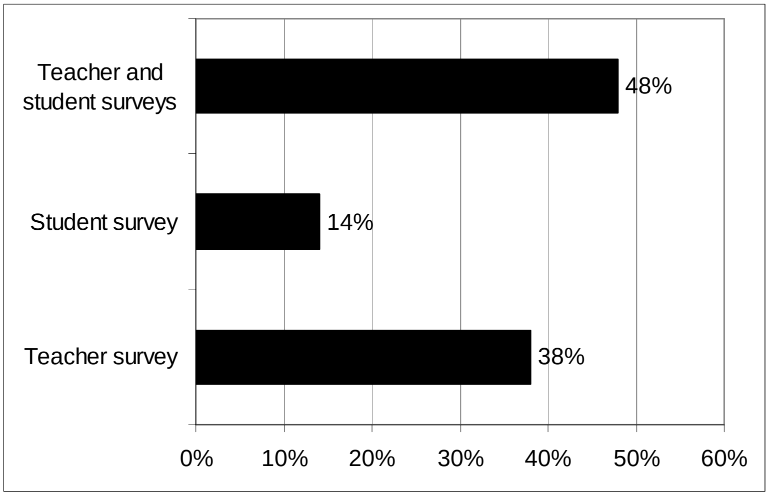


Figure 3. Percentage of studies using data from teacher surveys, student surveys, or both.

Indicators of Coverage

As noted above, we believe that the quality of the knowledge base on instruction depends in part on how well it covers major dimensions along which instruction varies. Reflecting the structure of the education system, instruction varies considerably across grade levels and subject areas. In designing the four large-scale surveys that are the focus of this research, NCES clearly intended to cover a wide range of grade levels and subject areas. Taken together, the four NCES surveys covered kindergarten through Grade 12, though each of the studies we reviewed focused on a subset of grade levels. Considered as a whole, the four surveys also focused on a wide range of subject areas, including English/language arts, mathematics, science, and social studies, though again, the individual studies we reviewed typically focused on only one or two subject areas.

We argued above that a knowledge base on instruction should have a critical mass of scholarship in each major curricular area. Figure 4 displays the percentage of studies in our database that focused on the major subjects of mathematics, science, English/language arts, and social studies. As the figure illustrates, the majority of the studies addressed a single academic subject, while only a handful addressed multiple subjects. Mathematics instruction has clearly received the greatest attention from researchers in this body of scholarship. Nearly half (46%) of the studies we analyzed focused on teaching mathematics. Science instruction was the focus of 13% of the studies, and science and math together accounted for 71%. Considering Figures 1 and 4 together, this lopsided distribution is not surprising. Studies using data from TIMSS focus exclusively on mathematics and science, many studies from NAEP focus on these subject areas, and studies using data from these two surveys make up 78% of all publications reviewed. We found that 15% of the studies focused on English/language arts and only 7% investigated social studies instruction.

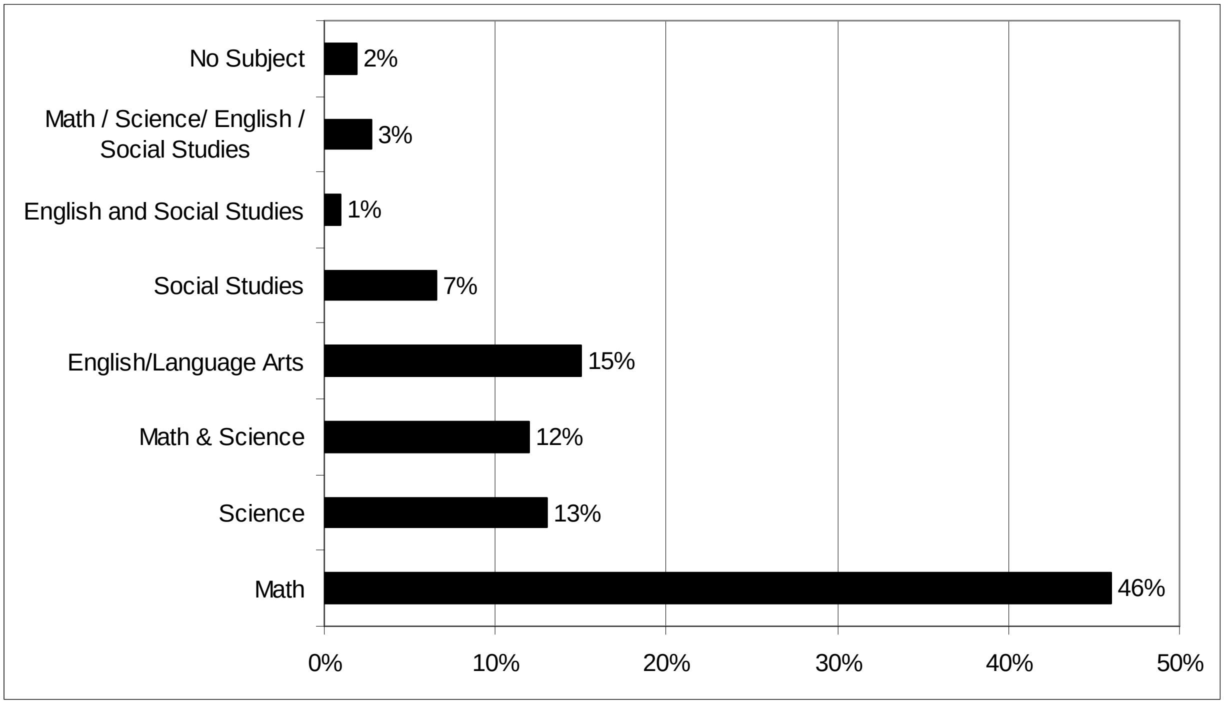
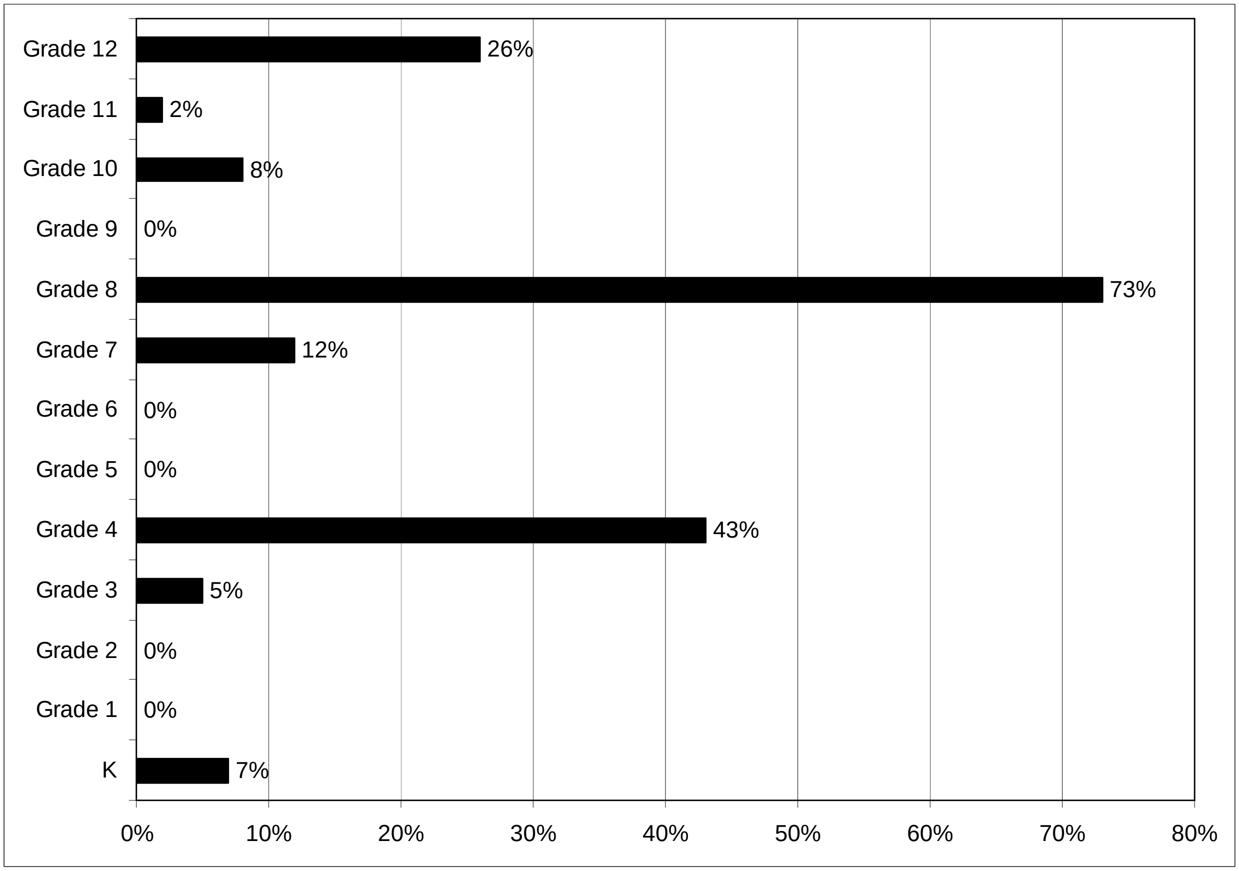


Figure 4. Percentage of studies that focused on various academic subjects.

Our second indicator of coverage was the extent to which different grade levels were represented in the body of research that was reviewed. As Figure 5 illustrates, the greatest concentrations of studies were in Grades 4 and 8, again reflecting the prevalence of studies that utilized data from TIMSS and NAEP. The relatively large percentage of studies examining 12th grade instruction was also driven by the NAEP survey, with a total of 21 studies using NAEP data from that grade level. We found that evidence on instruction in the early elementary grades was limited to a handful of studies using data from the ECLS survey. We found it quite striking that our review turned up no evidence on instruction in first grade, the pivotal grade in which most students learn to read, though we expect this to change in the next year or so as researchers begin to publish data from the second year of the ECLS survey. We also found that none of the studies reviewed investigated instruction in Grades 5 and 6, a period in which many students transition from elementary to middle school, or in Grade 9, another pivotal transition year.

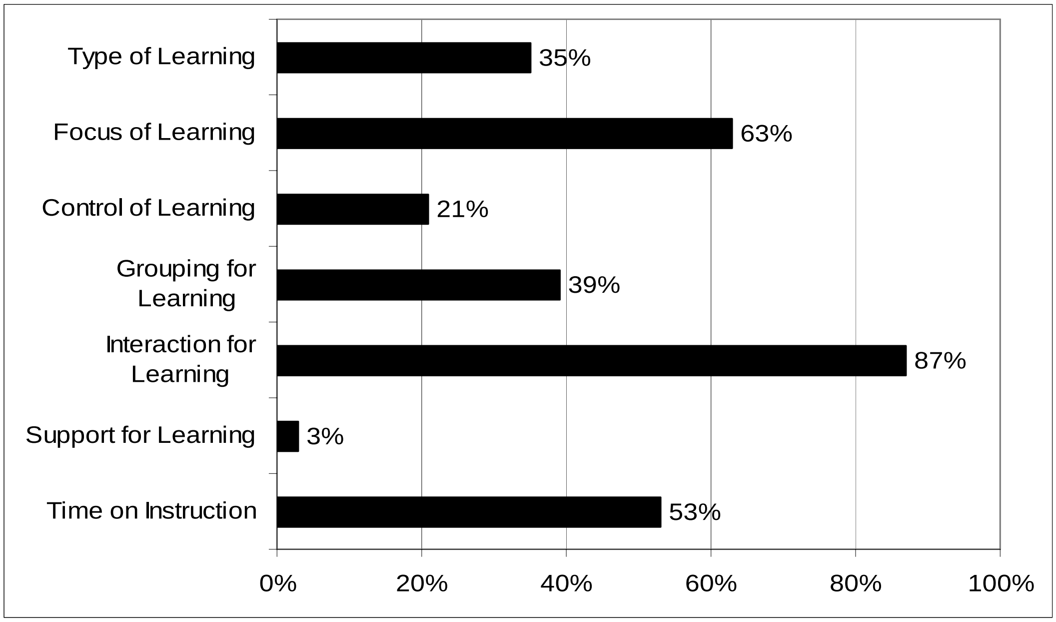
The third indicator of coverage we examined was dimension of instruction (Figure 6). We found that the most common dimension of instruction studied was *interactions for learning*. Recall that the interactions for learning dimension refers to the ways in which students interact with tools and materials and with other people in the classroom setting. A total of 87% of all studies that we reviewed focused on this dimension of instruction. As we discuss below, a substantial number of studies investigated students' use of manipulatives and tools such as calculators. Again, this focus is not terribly surprising given the large number of studies that examined mathematics instruction.



Note. Some studies focused on more than one grade.

Figure 5. Percentage of studies that focused on various grade levels.

Nearly two thirds of the studies reviewed examined the academic content of instruction, or what Reigeluth and Moore (1999) called the *focus of learning*. In the studies we reviewed, this dimension overwhelmingly referred to subject area and topic coverage. More than half of all studies reviewed examined *time on instruction* in some fashion. As we discuss below, many of the studies had a combined focus on instructional time and one of the other dimensions of instruction. For example, a substantial number looked at how much time (time on instruction) was spent on different curricular topics (focus of learning). A little more than one third (39%) looked at how students were grouped for instruction. However, as we indicate below, very few of the studies used inferential statistics, which was a requirement for studies to be included in the research summary. One third of the studies focused on the cognitive demands of instruction, referred to by Reigeluth and Moore (1999) as the *type of learning*. As we discuss below, a substantial proportion of these studies examined authentic instruction and problem solving in mathematics. Approximately one fifth of all studies examined the *control of learning*—that is, the degree to which instructional activities were controlled either by students or teachers. As Figure 6 illustrates, only 3% of the studies examined ways in which teachers provided *support for learning*.



Note. Some studies focus on more than one dimension of instruction.

Figure 6. Percentage of studies that focused on various dimensions of instruction.

Indicators of Scientific Quality

The vast majority of the studies we reviewed (72%) utilized simple descriptive statistics to investigate instruction (Figure 7). Many of these (58%) used multivariate analyses, but these analyses did not use procedures that allow for statistical inferences to be drawn. Twenty-eight percent of the studies used inferential statistics to investigate some aspect of instruction. These 27 studies are the focus of the research summary presented below. In nearly 40% of the studies we reviewed, instruction was analyzed as an independent variable, most often as a predictor of student achievement. In a slightly smaller percentage of studies, instruction was treated as a dependent variable, and the analytic strategy was to analyze factors that affected instruction. A small group of studies used multivariate analyses to examine instruction as both an independent and a dependent variable. In about one third of the studies, variation in patterns of instruction was simply described.

It is interesting to note that there was considerable variety in the analytic approaches of the 27 studies that used inferential statistics. The majority (21) used some form of linear model to test hypotheses about instruction. Fifteen used multilevel models that took into consideration hierarchical nesting in the data, while another six used linear regression models. Other analytic strategies that were used less frequently included structural equation modeling (two studies), simple correlations (two studies), and analysis of variance (one study). A minority of studies (nine) used longitudinal data to investigate instruction. Of these, eight examined the effect of instructional measures on longitudinal student achievement data—four examining achievement gains between two time points, and four predicting student achievement at one point in time while controlling for a measure of prior achievement (covariate adjustment model). We note, as have others (Rowan, Correnti, & Miller, 2002), that gains models and covariate adjustment models are among the weakest analytic strategies for predicting student achievement growth.

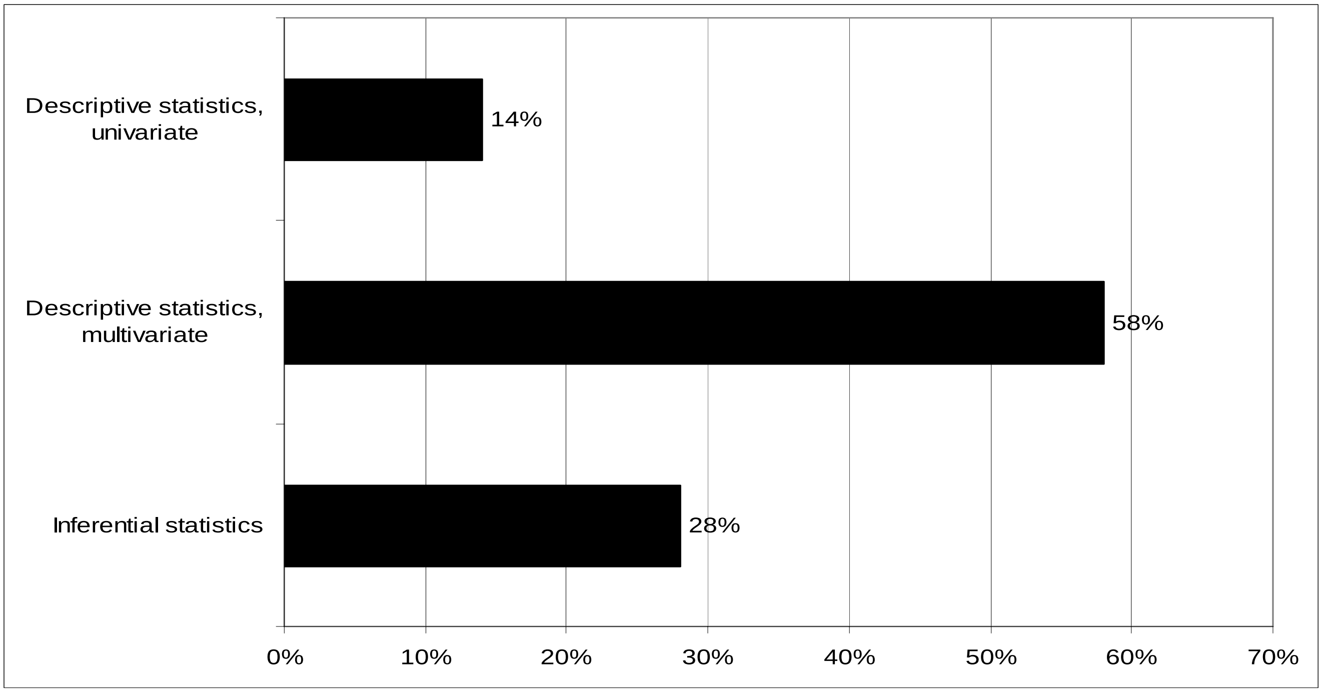


Figure 7. Percentage of studies using various analytic approaches.

The second indicator of scientific quality we examined was publication venue (Figure 8). A majority (58%) of the studies we reviewed were reports that were not subject to the same kind of peer-review process as papers published in refereed journals. We found that only about one quarter of the studies (23%) were published in refereed journals, while the remaining studies were either conference papers (12%) or books (2%).

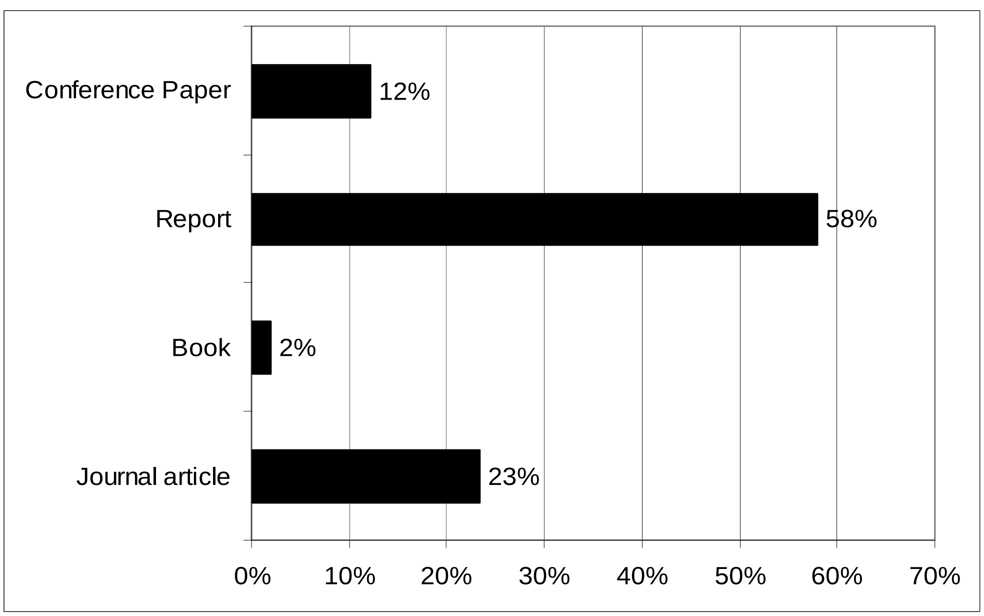


Figure 8. Percentage of studies published in various venues.

Narrative Summary of Results of NCES Studies Using Inferential Statistics

In this section, we present a narrative discussion of the 27 studies that used inferential statistics to study instruction. When relevant, we also mention the results of four additional studies that analyzed non-U.S. samples, noting the nation from which the sample was drawn. Our discussion is organized around the dimensions of instruction discussed previously. We note that there were no studies focusing on control for learning that used inferential statistics; therefore, our discussion focuses on the remaining six dimensions. For each dimension, we first summarize studies in which instruction was a dependent variable and then review studies in which instruction was an independent variable (usually predicting student achievement).

Type of Learning

Of the 27 studies we reviewed, 19 focused on the dimension of instruction Reigeluth and Moore (1999) called *type of learning*. This dimension refers to the cognitive goals of instruction and the cognitive activity demanded of students during instruction. For example, are students being asked to internalize the knowledge being presented? Are they being asked to apply knowledge to solve a problem? Are they being asked to engage in reasoning or analysis? Our review identified two categories of studies—(a) those that focused on authentic instruction or assessment and (b) those that focused on student thinking, reasoning, and problem solving.

Authentic instruction and assessment. In the 19 studies focusing on the type of learning implicit in instructional activities, authentic instruction and assessment received the greatest attention. Researchers have argued that instruction is authentic when the content students work on has value and meaning beyond the classroom. In authentic instruction, the real-world value and meaning of the knowledge are made clear to students, and instructional activities involve authentic social practice (Brown, Collins, & Duguid, 1989; Newmann & Wehlage, 1993).

Authentic instruction/assessment was found to occur more frequently in high–socioeconomic status (SES) middle schools (Wenglinsky, 2002) and in high schools with communal and student-centered organizational structures (Lee, Smith, & Croninger, 1997). Eighth-grade teachers were more likely to use authentic instruction when they received relevant professional development (Wenglinsky, 2002).

A number of studies in this group documented positive effects of authentic instruction on student achievement. For example, Wenglinsky (2004) found that schools in which higher percentages of eighth-grade students were encouraged to engage in real-world problem solving in mathematics tended to have higher levels of eighth-grade math achievement. Butty (2001) found that 12th-grade students who received what they called *reform instruction* in mathematics had higher achievement levels than students who received traditional instruction. Reform instruction included activities that illustrated the importance of mathematics to daily life, developed an awareness of the importance of mathematics academic fields, and showed the application of mathematics to business and industry. Similarly, outside the U.S. context, House (2004) showed that Japanese middle school students had higher mathematics achievement when their mathematics lessons were tied to “everyday life.” The results of studies assessing the effect of authentic instruction on student achievement were not unequivocal, however. For example, Wenglinsky (2002, 2003) found that the use of authentic assessment did not have a significant affect on achievement in eighth-grade mathematics or fourth-grade reading.

Two studies found that authentic instruction had a positive effect on students' enjoyment of learning. House (2002b) discovered that students found learning in science more enjoyable when instruction was tied to everyday life. House (2004) found that the same held true for Japanese students learning mathematics.

Student thinking, reasoning, and problem solving. A second set of studies examined student thinking, reasoning, and problem solving in instruction. Perkins (1992) argued that instruction that engages students in reasoning, thinking, and problem solving can lead to greater knowledge retention and understanding and more facile application of knowledge. Using NAEP data, Lubienski (2004) found that student engagement in collaborative problem solving was positively associated with mathematics achievement at both fourth and eighth grade, and that racial and socioeconomic achievement gaps were reduced when a measure of collaborative instruction was introduced into the statistical model.

Two studies found that exposure to instruction emphasizing thinking and reasoning varied across settings and subject areas. Wenglinsky (2003) found that fourth-grade students in high-SES schools received greater exposure to metacognition in their reading instruction than their peers in low-SES schools. Wenglinsky (2002) found that eighth-grade students in more affluent schools were more likely to be engaged in solving novel problems in mathematics. In contrast, Snow-Renner (2001) found that third- and fourth-grade mathematics classes tended to emphasize low-level skills and the traditional practice of computational skills.

Several studies found that instruction focused on thinking and reasoning had a positive effect on student achievement. Wenglinsky (2002) found that schools in which eighth-grade students had more opportunities to solve novel problems tended to have higher mathematics achievement. Wenglinsky (2003) discovered a positive effect of instruction involving metacognition on fourth-grade reading achievement. Gales and Yan (2001) found higher mathematics achievement among eighth-grade students who had more opportunities to discuss the reasoning behind their answers and consider several possible answers before determining their final answers. This same study also determined that students who spent more time working on problems for which an immediate correct answer was unavailable tended to have higher mathematics achievement. The results from this set of studies, like those from studies focusing on authentic instruction, were not unequivocal, however. Rowan, Chiang, and Miller (1997) found that teachers' emphasis on higher order thinking in 10^{th}-grade mathematics instruction did not have a significant effect on student achievement.

There was also some evidence that instruction in which students had *limited* opportunities to reason and engage in problem solving might have a *negative* effect on student achievement. Tomoff, Thompson, and Behrens (2000) found that seventh- and eighth-grade instruction that did not engage students in reasoning but instead had them practicing with rules and algorithms and engaging in drill and practice had a negative effect on their mathematics test scores.

Focus of Learning

Recall that the dimension of instruction called focus of learning refers to the relative emphasis teachers place on different curricular domains and topics. In the studies we examined, curricular and topic coverage were factors commonly targeted by researchers investigating the focus of learning in instruction. In some studies, curriculum coverage was operationalized quantitatively as the *amount* of material covered—for example, the percentage of the textbook completed or the number of topics taught. In other studies, the focus of learning was measured by noting *which* topics or knowledge domains were presented to students.

Four studies investigated the focus of learning as a dependent variable, looking at factors accounting for differences in students' opportunities to learn subject matter. Gau (1997) examined how the content of eighth-grade mathematics instruction varied with respect to average SES levels within schools. The author found that a smaller proportion of mathematics textbooks was covered by teachers in schools with more low-SES students than by teachers in schools with fewer such students. The author concluded that low-SES students' reduced exposure to mathematics content represented a diminished opportunity to learn. One of the questions Smerdon, Burkam, and Lee (1999) investigated was how the amount of time 10^{th}-grade teachers devoted to teaching science topics varied depending on the science course. Not surprisingly, they found that teachers spent more time on science topics in advanced science courses (honors biology and chemistry) than in lower level courses (basic biology, general biology, etc.). Using a similar standard for gauging students' learning opportunities in mathematics, Snow-Renner (2001) found that third-grade teachers covered fewer mathematics topics than fourth-grade teachers. Pong and Pallas (2001) investigated how eighth-grade teachers' mathematics topic coverage varied by class size and found topic exposure to be very similar in both of the class sizes studied.

A total of seven studies examined how the focus of learning in instruction affected student achievement. Five found a positive relationship between the amount of content covered in a subject area and student achievement in that subject area. For example, Gau (1997) discovered that eighth-grade students' math achievement was higher if their teachers covered larger percentages of mathematics textbooks. Tomoff et al. (2000) found a similar positive effect for textbook coverage in seventh- and eighth-grade mathematics. Snow-Renner (2001) determined that the number of mathematics topics covered was positively and significantly related to student mathematics achievement in the third and fourth grades. Pong and Pallas (2001) found that small class size had a positive effect on eighth-grade math achievement, though mathematics curriculum coverage did not mediate that relationship.

Three studies that examined the relationship between the focus of learning and student achievement looked beyond the quantity of content covered. Wenglinsky (2003) found that students whose teachers integrated reading and writing content tended to have higher reading comprehension achievement. Carbonaro and Gamoran (2002) determined that the topical focus in high school English classes, particularly a focus on *literature study* and *analytic writing*, had positive significant effects on growth in reading achievement. In general, measures of instructional content in this study had consistently stronger effects on achievement than measures of instructional practices. Lubienski (2004) found that a focus on non-number curricular strands was positively associated with fourth- and eighth-grade mathematics achievement. This study also found that this kind of curricular focus was associated with reductions in racial and SES achievement gaps.

Control of Learning

Four of the studies we reviewed examined the degree to which students or teachers controlled the pacing of learning. One looked at how often teachers controlled the pace of learning, while the remaining three examined how the control of learning affected achievement. Using ECLS data, Zeng and Zeng (2005) found that kindergarten teachers placed a high value on teacher control in the classroom. More than half of the teachers surveyed (53%) said they spent 2 hours or more on teacher-directed whole-class activities. In addition, vast majorities of teachers said it was very important or essential for students to "sit still and pay attention" and "follow directions."

The three studies examining the effect of the control of learning on achievement used markedly different approaches in operationalizing the independent variable. Walker (1999) found that students whose teachers used a more "student-centered" approach were more likely to correctly answer mathematics items measuring conceptual understanding. However, the items used by these investigators to measure student-centered learning—e.g., students' explaining their reasoning, students' working on "small investigations"—reflect the type of learning more than the control of learning according to our classification framework. Guthrie, Schafer, and Huang (2001) found that reading achievement was higher when students were permitted to take books from the library for enjoyment more frequently. Carbonaro and Gamoran (2002) examined how giving students greater "voice" in high school English classes affected their reading achievement. These investigators discovered that reading achievement was positively affected when students' demonstration of their understanding was a central part of instruction.

Grouping for Learning

Three of the studies we reviewed examined ways in which students were grouped for instruction and the impact that student grouping had on achievement. Smerdon et al. (1999) looked at different grouping configurations in 10^{th}-grade science classes and found whole-class instruction the most prevalent. Pong and Pallas (2001) similarly found whole-class instruction common in eighth-grade mathematics classrooms. This study also looked at school-level grouping practices as measured by class size. The authors found that class size was not a major determinant of teachers' instructional practices. This study also looked at how classroom grouping practices varied by class size and found that individualized instruction was no more prevalent in smaller classes than in larger classes.

Tomoff et al. (2000) examined the effect of student grouping on student achievement using TIMSS data. These researchers found that students' participation in group work was not a significant predictor of their mathematics achievement. However, analyzing science achievement among Hong Kong students, House (2000) found that seventh and eighth graders who worked together in pairs or small groups in science lessons had higher achievement.

Interactions for Learning

Of the studies we reviewed, 19 focused on ways in which students interacted with other students and teachers in the classroom or with instructional tools and artifacts such as computers, manipulatives, textbooks, and written assignments. The latter kind of interaction took a number of forms, such as hands-on learning, laboratory work in science, and the use of manipulatives in mathematics. Of the 19 studies that examined interactions for learning, 6 investigated instruction as an outcome variable, and 17 studied how interactions for learning affected the student outcomes of achievement and enjoyment of learning.

Of the studies examining interactions for learning as an outcome, three investigated factors that affect students' use of physical tools during instruction. Wenglinsky (2002) found that schools where teachers had received professional development in higher order thinking skills were more likely to have eighth-grade students engage in hands-on learning in mathematics. Lubienski (2004) documented a greater use of manipulatives in mathematics among fourth graders as compared to eighth graders. In their study of 10^{th}-grade science instruction, Smerdon et al. (1999) found that laboratory work was more common in chemistry and physical science classes than in other kinds of sciences classes. Tarr, Mittag, Uekawa, Kazuaki, and Lennex (2000) determined that calculator use in eighth-grade mathematics varied greatly from country to country. Wenglinsky (2003) found that among fourth graders, affluent students were more likely than their less affluent peers to read trade books in reading.

Two additional studies examined factors by which social interactions for learning varied. Lubienski (2004) found that collaborative instruction in mathematics is more common in fourth-grade than in eighth-grade classrooms. Pong and Pallas (2001) looked at whether instruction in small classes differed from that in larger classes and found that individual work with assistance from the teacher and class discussion were equally prevalent in all classrooms, regardless of size.

The studies that examined the effect of interactions for learning on student outcomes investigated three different aspects of this dimension of instruction—(a) students' interactions with physical tools or technology (15 studies), (b) students' interactions with written assignments and materials (10 studies), and (c) students' interactions with other people in the classroom (11 studies).

Among the studies investigating the effect of students' use of technology and physical tools, three looked at the effect of calculator use. Lubienski (2004) found the use of calculators was highly correlated with eighth-grade math achievement. This study also found substantial race-related gaps within each SES category, with both low- and high-SES White students having more calculator access than their respective Black and Hispanic peers. Byrnes (2003) found calculator use to be a positive predictor of 12^{th}-grade math achievement, Goertz (1994) found that calculator use explained considerable variation in 8^{th}-grade state NAEP math scores, and Tarr et al. (2000) found that 8^{th}-grade math achievement for U.S. students was positively associated with five different ways in which calculators were used. However, Tarr et al. (2000) also found a significant negative relationship between the frequency of student calculator use and student performance in eighth-grade math for Japanese students.

An additional two studies looked at the effect of computer use on student outcomes in the U.S. Franks, DeVaney, Weerakkody, & Katayama (1996) found that computer use did not have a significant effect on eighth-grade math achievement, and House (2002b) found that computer use in middle school science lessons was not significantly associated with student enjoyment of learning science. Similarly, House (2002a) found a significant negative relationship between the use of computers and mathematics achievement among seventh- and eighth-grade students in Taipei.

Two studies using U.S. samples examined the impact of the use of manipulatives or engagement in hands-on activities on student outcomes. Wenglinsky (2002) found that schools that placed a greater emphasis on hands-on learning in eighth-grade mathematics tended to have higher levels of mathematics achievement. Burkam, Lee, and Smerdon (1997) found that 10^{th}-grade students who participated in hands-on laboratory activities in science classes had significantly higher achievement, and this effect was especially pronounced for girls. In two studies of East Asian countries, House (2001, 2002a) found that middle school students who used objects from everyday life to solve problems had higher test scores in math. Two additional studies observed a positive relationship between the use of everyday objects in middle school science instruction and the enjoyment of science in the U.S. (House, 2002b, 2003).

The effects of students' interactions with written materials and assignments were assessed in a variety of ways. Wenglinsky (2004) found that at the school level, writing about math was negatively related to eighth-grade mathematics achievement. Rathbun and Hausken (2003) discovered that writing activity was not significantly related to reading gains in kindergarten. Wenglinsky (2004) found that time spent on test taking had a negative effect on eighth-grade mathematics achievement. Wenglinsky (2003) found that students' use of trade books was positively associated with fourth-grade reading comprehension, and House (2002b) determined that having middle school students look at science textbooks while teachers presented topics was associated with greater enjoyment of science. Time spent on homework was associated with higher achievement in eighth-grade math (Wenglinsky, 2004). Similarly, two studies of East Asian countries found that time spent on homework was associated with higher achievement in eighth-grade math (House, 2001) and seventh- and eighth-grade math (House, 2002a).

Two studies documented positive effects of students' work with peers on student outcomes. Guthrie et al. (2001) found that students' reading achievement in fourth grade was higher when they engaged in discussion with their classmates. House (2003) found that seventh- and eighth-grade students who engaged in cooperative learning were more likely to have a greater enjoyment of science. However, international studies that examined the effect of cooperative learning on student outcomes showed null or negative effects. House (2004) found that more frequent use of cooperative learning with students learning new mathematics topics was not significantly associated with mathematics achievement in Japan. In a study of Taipei students, House (2002a) documented a negative relationship between cooperative learning and mathematics achievement in seventh and eighth grade.

Teachers' interactions with students were shown to have a positive effect. For example, House (2002b) found that discussions between teachers and students had a positive effect on students' enjoyment of science in seventh and eighth grade. A number of international studies found that interactions between teachers and students had a similarly positive effect on seventh- and eighth-grade math achievement (House, 2001, 2002a) and seventh- and eighth-grade science achievement (House, 2000).

Time on Learning

As mentioned previously, Reigeluth and Moore's (1999) framework for comparing instructional strategies did not include a separate dimension for time spent on learning. It is true that a substantial number of studies that focused on the other dimensions of teaching included some element of time. As mentioned earlier, a number of studies found that the amount of time spent on homework had a positive impact on student achievement (Carbonaro & Gamoran, 2002; Gau, 1997; Wenglinsky, 2004). These studies could be classified under the Reigeluth and Moore dimensions *grouping for learning* (students typically work on homework by themselves) and *interactions for learning* (students interact with written materials/assignments). In contrast, 3 of the 27 studies that utilized inferential statistics contained analyses that focused exclusively on the amount of time spent on instruction. Gau (1997) found that, overall, lower SES students tended to receive lower amounts of instructional time in mathematics, though the relationship between school-level SES and instructional time varied by school sector. That study also found a negative effect of instructional time, as eighth-grade students whose mathematics classes met for longer periods tended to have lower mathematics achievement. Walston and West (2004) found that students who enrolled in full-day kindergarten made greater gains in reading and mathematics than students who enrolled in half-day kindergarten.

Conclusions

The research summary presented here provides a portrait of the contours of a generalizable knowledge base of quantitative evidence on K–12 instruction in the U.S that can be gleaned from national survey data. The portrait identifies areas in the literature where a critical mass of evidence has accumulated and also identifies gaps in the literature that could benefit from further attention from the research community. In this section, we summarize these peaks and valleys in the literature and offer some thoughts about potential implications of our research.

Before proceeding, we note three significant limitations of this research review. First, the portrait is not complete. The research reviewed here does not exhaust all large-scale survey research on instruction conducted in the last 20 years, and therefore we readily acknowledge that some of the gaps in the research we have identified may be filled by other studies that were not reviewed. Second, even though our narrative summary of research results focused only on studies utilizing inferential statistics, the majority of these utilized cross-sectional analyses. Moreover, the small number of studies conducting longitudinal analyses used among the weakest longitudinal statistical models. Consequently, there are substantial limitations in the causal inferences that can be drawn from these studies. Finally, in our view, the inferences that can be drawn from many of these studies are further limited by the ways in which instruction was conceptualized and operationalized. We, as others, view instruction as a complex social process that can be found in teachers' and students' *situated action* with academic content (Cohen & Ball, 1999; Lampert, 2001; Camburn & Barnes, 2004). In many studies reviewed here, however, instruction was conceptualized and measured in quite superficial ways. For example, the frequency of calculator use, the number of topics taught, and the number of minutes spent on instruction were some of the most common measures of instruction. It strikes us that research on instruction that gets beyond these kinds of surface features is likely to be more useful to policymakers and the practice community.

Our review revealed substantial evidence suggesting that learning opportunities are unevenly distributed among low- and high-income students. Among the 27 studies reviewed for the research summary, 9 examined instruction as an outcome variable. These nine studies provided useful insight into how student learning opportunities are distributed across various student groups. Most striking was a repeated finding that low-SES students had diminished learning opportunities compared to their more affluent peers. Our summary found that lower SES students received less time on instruction overall and were exposed to a smaller proportion of mathematics textbooks than more affluent students. The research summary also revealed that lower SES students were less likely to (a) be exposed to authentic instruction, (b) engage in meta-cognition and problem solving, and (c) read trade books. Lubienski's (2004) finding that particular instructional experiences were associated with reductions in racial and SES disparities in mathematics achievement suggests that SES-related disparities in instruction can be consequential.

The research summary also showed that student learning opportunities can vary substantially from one grade and one class to the next. For example Snow-Renner (2001) found that fewer mathematics topics were taught in third grade than in fourth grade, and Lubienski (2004) found that fourth graders were more likely than eighth graders to use manipulatives in mathematics. Smerdon et al. (1999) found that high school science courses differed substantially in the amount of time devoted to teaching science topics and in the amount of lab work that took place in class. While the variation in learning opportunities associated with SES has clear negative implications, it is not entirely clear to us whether such between-grade and between-course variation is problematic or whether it reflects equitable and educationally effective variation within the education system.

Our review also identified six instructional activities that had a positive effect on student achievement. We considered an activity to have a positive effect if four or more studies investigated the impact of the activity and the majority of these found positive effects on achievement.[5] The research summary demonstrated that the following instructional activities met these criteria: authentic instruction, instruction involving thinking or reasoning, the amount of academic content covered, the use of calculators, the use of manipulatives, and teachers' interactions with students. As discussed above, evidence in support of the effect of these activities on achievement was not unequivocal. However, the preponderance of evidence on these strategies suggests that they can have a positive impact on students.

We also found either negative or mixed results for two instructional activities. While a number of studies showed that students' interactions with written materials and assignments (e.g., homework, in-class written assignments) had a positive impact on student achievement,

[5] We found fewer than four studies using inferential statistics for three dimensions of instruction—control of learning, grouping for learning, and time on learning—and therefore we were unable to make summary judgments about the effect of instructional activities on student achievement in those areas.

others found either a negative effect or no effect at all. Similarly, two studies examining the effect of computer use on achievement found no effects, while another study observed a significant negative effect on achievement.

The research summary identified significant gaps in the research. Our results indicate a fairly uneven distribution of studies across dimensions of instruction and indicators of scientific quality. We found no studies that examined instruction during first grade and at the important transition points of fifth, sixth, and ninth grades. Moreover, when you consider the grade levels covered by the studies reviewed here (see Figure 5)—as well as the NRC research summaries and the NRP report discussed earlier, which focused on instruction in the elementary grades—it becomes clear that there is a marked lack of empirical evidence on instruction in the first 3 years of high school.

Generalizable evidence on instruction from large-scale national surveys also appears to be somewhat dated, with more than half the studies using data that are more than a decade old. As we noted earlier, this tendency to use older data raises a question in our minds about whether this evidence base can adequately provide researchers and policymakers with information that is relevant to current educational and policy conditions.

Our summary also indicates that evidence on instruction from large-scale national surveys focuses disproportionately on the fields of mathematics and science. As we noted above, this result is not entirely surprising given the focus of TIMSS and NAEP and the widespread use of data from these surveys by the research community. Again, when considering a broader array of potential sources of generalizable evidence on instruction, the NRP report and the NRC report on preventing reading difficulties somewhat counterbalance the lopsided focus of studies using large-scale national surveys. However, our review still suggests a relative paucity of research on instruction in English/language arts and social studies, with the latter academic discipline having by far the smallest amount of generalizable quantitative evidence on instruction available.

Finally, our review of research on instruction raises a question about the scientific quality of this body of research. Less than one third of the studies we reviewed were published in peer-refereed journals. In addition, we could only find 27 studies that used inferential statistics to test research questions about instruction. Indeed, a slim majority of the studies we reviewed were published reports that used descriptive statistics to analyze instruction. This result raises doubts in our minds about the validity and robustness of knowledge that can be gleaned from this collection of studies.

A number of implications for research on instruction logically follow from these results. More research is needed on what appear to be substantial disparities between low- and high-income students' instructional experiences. There is also a need for research on instruction at key transition points, particularly Grades 5, 6, and 9. Research indicates that students often have difficulty negotiating these transitions (see, for example, Roderick & Camburn, 1999), and gaining a better understanding of students' classroom experiences in these grades may shed light on these difficulties. In addition, as mentioned above, our review reinforces the point that more research is needed on instruction in high school, particularly in Grades 9–11. This review also indicates a need for more research on instruction in English/language arts and social studies. Finally, we identify a need for the education research community to engage in more scientifically rigorous inquiries on instruction. In particular, our review suggests a need for studies that more rigorously test research questions about factors that affect instruction and the ways in which different aspects of instruction affect student achievement.

In our view, filling these gaps in the literature and bringing an increased rigor to the study of instruction are vitally important to the field. As we and many others have argued, instruction is a fundamental determinant of student learning, and consequently, the question of how instruction is practiced, to whom it is delivered, and how it affects student learning deserves serious attention from the education research community. We hope that this research summary will serve as a useful resource for future inquiries.

Cited References

Ball, D. L., & Rowan, B. (2004). Introduction: Measuring instruction. *Elementary School Journal, 105*(1), 3–10.

Brown, J., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher, 18*(1), 32–42.

Burkam, D. T., Lee, V. E., & Smerdon, B. A. (1997). Gender and science learning early in high school: Subject matter and laboratory experiences. *American Educational Research Journal, 34*(2), 297–331.

Butty, J.-A. L. M. (2001). Teacher instruction, student attitudes, and mathematics performance among 10th and 12th grade black and Hispanic students. *Journal of Negro Education, 70*(1/2), 19–37.

Byrnes, J. P. (2003). Factors predictive of mathematics achievement in white, black, and Hispanic 12th graders. *Journal of Educational Psychology, 95*(2), 316–326.

Camburn, E., & Barnes, C. (2004). Assessing the validity of a language arts instruction log through triangulation. *Elementary School Journal*, *105*, 49–74.

Carbonaro, W. J., & Gamoran, A. (2002). The production of achievement inequality in high school English. *American Educational Research Journal*, *39*(4), 801–827.

Cohen, D. K., & Ball, D. L. (1999). *Instruction, capacity and improvement* (CPRE Research Report Series RR-43). Philadelphia: University of Pennsylvania, Consortium for Policy Research in Education.

Cuban, L. (1993). *How teachers taught: Constancy and change in American classrooms, 1890–1990*. New York: Teachers College Press.

Franks, M. E., DeVaney, T. A., Weerakkody, G. J., & Katayama, A. D. (1996, April). *An analysis of NAEP trial state assessment data concerning the effects of computers on mathematics achievement*. Paper presented at the annual meeting of the American Educational Research Association. New York.

Gales, M. J., & Yan, W. (2001, April). *Relationship between constructivist teacher beliefs and instructional practices to students' mathematical achievement: Evidence from TIMMS.* Paper presented at the annual meeting of the American Educational Research Association, Seattle, WA.

Gau, S.-J. (1997, April). *The distribution and the effects of opportunity to learn on mathematics achievement.* Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Goertz, M. E. (1994). *Opportunity to learn: Instructional practices in eighth-grade mathematics: Data from the 1990 NAEP trial state assessment* (CPRE Research Report Series Report #32). Brunswick, NJ: Consortium for Policy Research in Education.

Guthrie, J. T., Schafer, W. D., & Huang, C.-W. (2001). Benefits of opportunity to read and balanced instruction on the NAEP. *Journal of Educational Research, 94*(3), 145–162.

House, J. D. (2000). Relationship between instructional activities and science achievement of adolescent students in Hong Kong: Findings from the Third International Mathematics and Science Study (TIMSS). *International Journal of Instructional Media, 27*(3), 275–288.

House, J. D. (2001). Relationship between instructional activities and mathematics achievement of adolescent students in Japan: Findings from the Third International Mathematics and Science Study (TIMSS). *International Journal of Instructional Media, 28*(1), 93–105.

House, J. D. (2002a). Instructional practice and mathematics achievement of adolescent students in Chinese Taipei: Results from the TIMSS 1999 assessment. *Child Study Journal, 32*(3), 157–178.

House, J. D. (2002b). The motivational effects of specific teaching activities and computer use for science learning: Findings from the Third International Mathematics and Science Study (TIMSS). *International Journal of Instructional Media, 29*(4), 423–439.

House, J. D. (2003). Instructional activities and interest in science learning for adolescent students in Japan and United States: Findings from the Third International Mathematics and Science Study (TIMSS). *International Journal of Instructional Media, 31*(2), 429–442.

House, J. D. (2004). The effects of homework activities and teaching strategies for new mathematics of adolescent students in Japan: Results from the TIMSS 1999 assessment. *International Journal of Instructional Media, 31*(2), 199–120.

Lampert, M. (2001). *Teaching problems and the problems of teaching*. New Haven, CT: Yale University Press.

Lee, V. E., Smith, J. B., & Croninger, R. G. (1997). How high school organization influences the equitable distribution of learning in mathematics and science. *Sociology of Education, 70*(2), 128–150.

Lubienski, S. T. (2004, April). *Instruction, achievement and equity: Intersections of race/ethnicity and SES in NAEP data*. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.

National Institute of Child Health and Human Development. (2000). *Report of the National Reading Panel. 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* (NIH Publication No. 00-4754). Washington, DC: U.S. Government Printing Office.

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

Newmann, F., & Wehlage, G. (1993). Five standards of authentic instruction. *Educational Leadership, 50*(7), 8–12.

Perkins, D. (1992). *Smart schools: From training memories to educating minds.* New York: Free Press.

Pong, S.-l., & Pallas, A. (2001). Class size and eighth-grade math achievement in the United States and abroad. *Educational Evaluation and Policy Analysis, 23*(3), 251–273.

Rathbun, A. H., & Hausken, E. G. (2003). *Reading and writing instruction in kindergarten: How often and who receives it?* Chicago: American Educational Research Association.

Rathbun, A. H., Walston, J. T., & Hausken, E. G. (2000, April). *Kindergarten teachers' use of developmentally appropriate practices: Results from the Early Childhood Longitudinal Study, kindergarten class of 1998–1999.* Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Reigeluth, C. M., & Moore, J. (1999). Cognitive education and the cognitive domain. In C. Reigeluth (Ed.), *Instructional-design theories and models: Vol. II. A new paradigm of instructional theory* (pp. 51–68). Mahwah, NJ: Lawrence Erlbaum.

Roderick, M., & Camburn, E. (1999). Risk and recovery from course failure in the early years of high school. *American Educational Research Journal, 36*(2), 303–343.

Rowan, B., Chiang, F. S., & Miller, R. J. (1997). Using research on employees' performance to study the effects of teachers on students' achievement. *Sociology of Education, 70*(4), 256–284.

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*(8), 1525–1567.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference.* Boston: Houghton Mifflin.

Smerdon, B. A., Burkam, D. T., & Lee, V. E. (1999). Access to constructivist and didactic teaching: Who gets it? Where is it practiced? *Teachers College Record, 101*(1), 5–34.

Snow, C. E., Burns, M. S., & Griffin, P. (Eds.). (1998). *Preventing reading difficulties in young children*. Washington DC: National Academy Press.

Snow-Renner, R. (2001, April). *What is the promise of large-scale classroom practice measures for informing us about equity in student opportunities-to-learn? An example using the Colorado TIMSS*. Paper presented at the annual meeting of the American Educational Research Association, Seattle, WA.

Tarr, J. E., Mittag, K. C., Uekawa, Kazuaki, & Lennex, L. (2000). A comparison of calculator use in eighth-grade mathematics classrooms in the United States, Japan, and Portugal: Results from the Third International Mathematics and Science Study. *School Science and Mathematics, 100*(3), 139–150.

Tomoff, J., Thompson, M., & Behrens, J. (2000). *Measuring NCTM-recommended practices and student achievement with TIMSS*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Walston, J., & West, J. (2004). *Full-day and half-day kindergarten in the United States: Findings from the Early Childhood Longitudinal Study, kindergarten class of 1998–99* (NCES 2004-078). Washington, DC: National Center for Education Statistics.

Wenglinsky, H. (2002). The link between teacher classroom practices and student academic performance. *Education Policy Analysis Archives, 10*(2). Retrieved January 23, 2008, from http://epaa.asu.edu/epaa/v10n12/

Wenglinsky, H. (2003). Using large-scale research to gauge the impact of instructional practices on student reading comprehension: An exploratory study. *Education Policy Analysis Archives, 11*(9). Retrieved January 23, 2008, from http://epaa.asu.edu/epaa/v11n19/

Wenglinsky, H. (2004). The link between instructional practice and the racial gap in middle schools. *Research in Middle Level Education, 28*(1). Retrieved January 23, 2008, from http://www.nmsa.org/portals/0/pdf/publications/RMLE/rmle_vol28_no1_article1.pdf

Studies Using NCES Survey Data That Were Reviewed

TIMSS

Beaton, A. E., Mullis, I., V. S., Martin, M. O., Gonzalez, E. J., Kelly, D. L., & Smith, T. A. (1996a). *Mathematics achievement in the middle school years: IEA's Third International Mathematics and Science Study*. Chestnut Hill, MA: Boston College, TIMSS International Study Center.

Beaton, A. E., Mullis, I., V. S., Martin, M. O., Gonzalez, E. J., Kelly, D. L., & Smith, T. A. (1996b). *Science achievement in the middle school years: IEA's Third International Mathematics and Science Study*. Chestnut Hill, MA: Boston College, TIMSS International Study Center.

Cwikla, J. (2001). *TIMSS-R 1999: An analysis of the Delaware science coalition data* (TIMSS-R report). Newark: University of Delaware, Delaware Education Research & Development Center.

Cwikla, J. (2002). *Top TIMSS-R mathematics performers: What are they doing differently?* (TIMSS-R Report 2002). Wilmington: Delaware Foundation for Science & Math Education.

Gales, M. J., & Yan, W. (2001, April). *Relationship between constructivist teacher beliefs and instructional practices to students' mathematical achievement: Evidence from TIMMS.* Paper presented at the annual meeting of the American Educational Research Association, Seattle, WA.

House, J. D. (2002b). The motivational effects of specific teaching activities and computer use for science learning: Findings from the Third International Mathematics and Science Study (TIMSS). *International Journal of Instructional Media, 29*(4), 423–439.

House, J. D. (2003). Instructional activities and interest in science learning for adolescent students in Japan and United States: Findings from the Third International Mathematics and Science Study (TIMSS). *International Journal of Instructional Media, 31*(2), 429–442.

Kimmelman, P., Kroeze, D., Schmidt, W., van der Ploeg, A., McNeely, M., & Tan, A. (1999). *A first look at what we can learn from high performing school districts: An analysis of TIMSS data from the first in the world consortium*. Washington, DC: U.S. Department of Education, National Institute on Student Achievement, Curriculum, and Assessment.

Leung, F. K. S. (2002). Behind the high achievement of East Asian students. *Educational Research and Evaluation: An International Journal on Theory and Practice, 8*(1), 87–108.

Martin, M. O., Mullis, I., V. S., Beaton, A. E., Gonzalez, E. J., Smith, T. A., & Kelly, D. L. (1997). *Science achievement in the primary school years: IEA's Third International*

Mathematics and Science Study. Chestnut Hill, MA: Boston College, TIMSS International Study Center.

Martin, M. O., Mullis, I., V. S., Beaton, A. E., Gonzalez, E. J., Smith, T. A., & Kelly, D. L. (1998). *Science achievement in Missouri and Oregon in an international context: 1997 TIMSS benchmarking*. Chestnut Hill, MA: Boston College, Center for the Study of Testing, Evaluation, and Educational Policy.

Martin, M. O., Mullis, I., V. S., Gonzalez, E. J., & Chrostowski, S. J. (2004). *TIMSS 2003 international science report: Findings from IEA's Trends in International Mathematics and Science Study at the fourth and eighth grades.* Chestnut Hill, MA: Boston College, TIMSS & PIRLS International Study Center.

Martin, M. O., Mullis, I., V. S., Gonzalez, E. J., Gregory, K. D., Smith, T., A., Chrostowski, S. J., et al. (2000). *TIMSS 1999 international science report: Findings from IEA's repeat of the Third International Mathematics and Science Study at the eighth grade*. Chestnut Hill, MA: Boston College, TIMSS & PIRLS International Study Center.

Martin, M. O., Mullis, I., V.S., Gonzalez, E. J., O'Connor, K. M., Chrostowski, S. J., Gregory, K. D., et al. (2001). *Science benchmarking report: TIMSS 1999 — Eighth grade: Achievement for U.S. states and districts in an international context.* Chestnut Hill, MA: Boston College, International Study Center.

Masini, B. E. (2001, April). *Race differences in exposure to algebra and geometry among U.S. eighth-grade students.* Paper presented at the annual meeting of the American Educational Research Association, Seattle, WA.

Mullis, I., V. S., Martin, M. O., Beaton, A. E., Gonzalez, E. J., Kelly, D. L., & Smith, T. A. (1997). *Mathematics achievement in the primary school years: IEA's Third International Mathematics and Science Study*. Chestnut Hill, MA: Boston College, TIMSS International Study Center.

Mullis, I., V. S., Martin, M. O., Beaton, A. E., Gonzalez, E. J., Kelly, D. L., & Smith, T., A. (1998). *Mathematics achievement in Missouri and Oregon in an international context: 1997 TIMSS benchmarking*. Chestnut Hill, MA: Boston College, TIMSS International Study Center.

Mullis, I., V. S., Martin, M. O., Gonzalez, E. J., & Chrostowski, S. J. (2004). *TIMSS 2003 International mathematics report: Findings from IEA's trends in International Mathematics and Science Study at the fourth and eighth grades*. Chestnut Hill, MA: Boston College, TIMSS & PIRLS International Study Center.

Mullis, I. V. S., Martin, M. O., Gonzalez, E. J., Gregory, K. D., Garden, R. A., O'Connor, K. M., et al. (2000). *TIMSS 1999 International mathematics report: Findings from IEA's repeat of the Third International Mathematics and Science Study at the eighth grade*. Chestnut Hill, MA: Boston College, International Study Center.

Mullis, I., V. S., Martin, M. O., Gonzalez, E. J., O'Connor, K. M., Chrostowski, S. J., Gregory, K. D., et al. (2001). *Mathematics benchmarking report: TIMSS 1999—eighth grade: Achievement results for U.S. states and districts in an international context.* Chestnut Hill, MA: Boston College, International Study Center.

National Education Goals Panel. (2000). *Minnesota & TIMSS: Exploring high achievement in eighth grade science: Lessons from the states*. Washington, DC: Author.

Pong, S.-l., & Pallas, A. (2001). Class size and eighth-grade math achievement in the United States and abroad. *Educational Evaluation and Policy Analysis, 23*(3), 251–273.

Rowan, B., Chiang, F. S., & Miller, R. J. (1997). Using research on employees' performance to study the effects of teachers on students' achievement. *Sociology of Education, 70*(4), 256–284.

Schmidt, W. H., McKnight, C. C., Raizen, S. A., & Third International Mathematics and Science Study. (1997). *A splintered vision: An investigation of U.S. science and mathematics education*. Dordrecht, Holland: Kluwer Academic Publishers.

Snow-Renner, R. (2001, April). *What is the promise of large-scale classroom practice measures for informing us about equity in student opportunities-to-learn? An example using the Colorado TIMSS*. Paper presented at the annual meeting of the American Educational Research Association, Seattle, WA.

Tananis, C. A., Chrostowski, S. J., Bunt, N. R., Seeley, M. M., & Tamler, L. (2002). *TIMSS 1999 southwest Pennsylvania: Third International Mathematics and Science Study, eighth grade: Regional benchmarking report, achievement for a workforce region in a national and international context*. Chestnut Hill, MA: Boston College, International Study Center.

Tarr, J. E., Mittag, K. C., Uekawa, Kazuaki, & Lennex, L. (2000). A comparison of calculator use in eighth-grade mathematics classrooms in the United States, Japan, and Portugal: Results from the Third International Mathematics and Science Study. *School Science and Mathematics, 100*(3), 139–150.

Tomoff, J., Thompson, M., & Behrens, J. (2000). *Measuring NCTM-recommended practices and student achievement with TIMSS*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Walker, C. M. (1999, April). *The effect of different pedagogical approaches on mathematics students' achievement*. Paper presented at the annual meeting of the American Educational Research Association, Montreal, Quebec, Canada.

Williams, T., Levine, D., Jocelyn, L., Butler, P., Heid, C., & Haynes, J. (2000). *Mathematics and science in the eighth grade: Findings from the Third International Mathematics and Science Study*. Washington DC: National Center for Education Statistics.

NAEP

Allen, R., Bettis, N., Kurfman, D., MacDonald, W., Mullis, I. V. S., & Salter, C. (1990). *The geography learning of high-school seniors*. Washington, DC: Office of Educational Research and Improvement.

Anderson, L., Jenkins, L. B., Leming, J., MacDonald, W. B., Mullis, I. V. S., Turner, M. J., et al. (1990). *The civics report card: Trends in achievement from 1976 to 1988 at ages 13 and 17; Achievement in 1988 at grades 4, 8, and 12*. Princeton, NJ: Educational Testing Service.

Applebee, A. N., Langer, J. A., Jerkins, L. B., Mullis, I. V. S. & Foertsch, M. A. (1990). *Learning to write in our nation's schools: Instruction and achievement in 1988 at grades 4, 8, and 12*. Princeton, NJ: Educational Testing Service.

Applebee, A. N., Langer, J. A., & Mullis, I. V. S. (1988). *Who reads best? Factors related to reading achievement in grades 3, 7, and 11* (Report No. 17-R-01). Princeton, NJ: Educational Testing Service.

Barton, P. E., & Coley, R. J. (1994). *Testing in America's schools. Policy information report*. Princeton, NJ: Educational Testing Service.

Beatty, A. S., Reese, C. M., Persky, H. R., & Carr, P. (1996). *NAEP 1994 U.S. history report card: Findings from the National Assessment of Educational Progress*. Princeton, NJ: Educational Testing Service.

Blank, R. K., & Gruebel, D. (1995). *State indicators of science and mathematics education 1995: State-by-state trends and new indicators from the 1993–94 school year*. Washington, DC: Council of Chief State School Officers.

Blank, R. K., & Langesen, D. (1999). *State indicators of science and mathematics education, 1999: State-by-state trends and new indicators from the 1997–98 school year*. Washington, DC: Council of Chief State School Officers.

Braswell, J. S., Lutkus, A. D., Grigg, W. S., Santapau, S. L., Tay-Lim, B., & Johnson, M. (2001). *The nation's report card: Mathematics, 2000*. Washington, DC: National Center for Education Statistics.

Byrnes, J. P. (2003). Factors predictive of mathematics achievement in white, black, and Hispanic 12th graders. *Journal of Educational Psychology, 95*(2), 316–326.

Coley, R. J., & Coleman, A. B. (2004). *The fourth-grade reading classroom. Policy information report*. Princeton, NJ: Educational Testing Service.

Dossey, J. A., Mullis, I. V. S., Gorman, S., & Latham, A. S. (1994). *How school mathematics functions: Perspectives from the NAEP 1990 and 1992 assessments*. Washington, DC: U.S. Government Printing Office.

Educational Testing Service. (1994). *Data compendium for the NAEP 1992 reading assessment of the nation and the states*. Washington, DC: Author.

Foertsch, M. A. (1992). *Reading in and out of school: Factors influencing the literacy achievement of American students in grades 4, 8, and 12, in 1988 and 1990*. Washington, DC: U.S. Government Printing Office.

Franks, M. E., DeVaney, T. A., Weerakkody, G. J., & Katayama, A. D. (1996, April). *An analysis of NAEP trial state assessment data concerning the effects of computers on mathematics achievement*. Paper presented at the annual meeting of the American Educational Research Association. New York.

Goertz, M. E. (1994). *Opportunity to learn: Instructional practices in eighth-grade mathematics: Data from the 1990 NAEP trial state assessment* (CPRE Research Report Series Report #32). Brunswick, NJ: Consortium for Policy Research in Education.

Greenberg, E. (1995). More metro than nonmetro students have access to computers, but their rates of usage are similar. *Rural Development Perspectives, 10*(3), 61–64.

Grouws, D. A., Smith, M. S., & Sztajn, P. (2004). The preparation and teaching practices of United States mathematics teachers: Grades 4 and 8. In P. Kloosterman & F. K. Lester (Eds.), *Results and interpretations of the 1990 through 2000 mathematics assessments of the National Assessment of Educational Progress* (pp. 221–267). Reston, VA: National Council of Teachers of Mathematics.

Guthrie, J. T., Schafer, W. D., & Huang, C.-W. (2001). Benefits of opportunity to read and balanced instruction on the NAEP. *Journal of Educational Research, 94*(3), 145–162.

Hammack, D. C., Hartoonian, M., Howe, J., Jenkins, L. B., Levstik, L. S., MacDonald, W. B., et al. (1990). *The U.S. history report card: The achievement of fourth-, eighth-, and twelfth-grade students in 1988 and trends from 1986 to 1988 in the factual knowledge of high-school juniors*. Princeton, NJ: Educational Testing Service.

Hawkins, E. F., Stancavage, F. B., & Dossey, J. A. (1998). *School policies and practices affecting instruction in mathematics: Findings from the National Assessment of Educational Progress*. Princeton, NJ: Educational Testing Service.

Kloosterman, P., Warfield, J., Wearne, D., Koc, Y., Martin, W. G., & Strutchens, M. (2004). Fourth-grade students' knowledge of mathematics and perceptions of learning mathematics. In P. Kloosterman & F. K. Lester (Eds.), *Results and interpretations of the 1990 through 2000 mathematics assessments of the National Assessment of Educational Progress* (pp. 71–103). Reston, VA: National Council of Teachers of Mathematics.

Langer, J. A., Applebee, A. N., Mullis, I. V. S., & Foertsch, M. A. (1990). *Learning to read in our nation's schools: Instruction and achievement in 1988 at grades 4, 8, and 12*. Princeton, NJ: Educational Testing Service.

Lindquist, M. M. (1997). NAEP Findings regarding the preparation and classroom practices of mathematics teachers. In P. A. Kenney & E. A. Silver (Eds.), *Results from the sixth mathematics assessment of the National Assessment of Educational Progress* (pp. 61–86). Reston, VA: National Council of Teachers of Mathematics.

Lubienski, S. T. (2002). A closer look at black-white mathematics gaps: Intersections of race and SES in NAEP achievement and instructional practices data. *Journal of Negro Education, 71*(4), 269–287.

Lubienski, S. T. (2004, April). *Instruction, achievement and equity: Intersections of race/ethnicity and SES in NAEP data*. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.

Lubienski, S. T., & Shelley, M. C., II. (2003, April). *A closer look at U.S. mathematics instruction and achievement: Examinations of race and SES in a decade of NAEP data*. Paper presented at the annual meeting of the American Educational Research Association, Chicago.

Mangels, K. (1993, April). *A study of mathematics education in the Lutheran Church–Missouri Synod schools using background questions from the 1990 National Assessment of Educational Progress*. Paper presented at the annual meeting of the American Educational Research Association, Atlanta, GA.

Mitchell, J. H., Hawkins, E. F., Jakwerth, P. M., Stancavage, F. B, & Dossey, J. A. (1999). *Student work and teacher practices in mathematics*. Washington, DC: National Center for Education Statistics.

Mullis, I. V. S., Campbell, J. R., & Farstrup, A. E. (1993). *Executive summary of the NAEP 1992 reading report card for the nation and the states: Data from the national and trial state assessments*. Washington, DC: U.S. Government Printing Office.

Mullis, I. V. S. (Ed.) (1994). *America's mathematics problem: Raising student achievement. A Synthesis of Findings from NAEP's 1992 Mathematics Assessment.* (NAEP-23-FR-03; NCES-94-703). Princeton, NJ: National Assessment of Educational Progress.

Mullis, I. V. S., Jenkins, F., & Johnson, E. G. (1994). *Effective schools in mathematics: Perspectives from the NAEP 1992 assessment*. Washington, DC: National Center for Education Statistics.

National Center for Education Statistics. (1992). *NAEPfacts: Calculators and computers.* Washington, DC: Author.

National Center for Education Statistics. (1993). *NAEP 1992: Reading state report for Colorado: The trial state assessment at grade 4*. Washington, DC: Author.

National Center for Education Statistics. (2002). *The nation's report card: Geography highlights, 2001*. Washington, DC: Author.

O'Sullivan, C. Y., & Weiss, A. R. (1999). *Student work and teacher practices in science: A report on what students know and can do*. Washington, DC: U.S. Government Printing Office.

O'Sullivan, C. Y., Weiss, A. R., & Askew, J. M. (1998). *Students learning science: A report on policies and practices in U.S. schools*. Washington, DC: U.S. Government Printing Office.

Shaughnessy, J. M., & Zawojewski, J. S. (1999). Secondary students' performance on data and chance in the 1996 NAEP. *Mathematics Teacher, 92*(8), 713–718.

Sowder, J. T., Wearne, D., Martin, W. G., & Strutchens, M. (2004). What do 8^{th}-grade students know about mathematics? Changes over a decade. In P. Kloosterman & F. K. Lester (Eds.), *Results and interpretations of the 1990 through 2000 mathematics assessments of the National Assessment of Educational Progress* (pp. 105–143). Reston, VA: National Council of Teachers of Mathematics.

Swanson, C. B., Plank, S. B., & Hewes, G. M. (2003). *From national movement to local action: The status of standards-based science instruction in middle school classrooms* (CRESPAR-RN-64). Baltimore, MD: Center for Research on the Education of Students Placed at Risk.

Unger, J., & Fleischman, S. (2004). Is process writing the 'write stuff'? *Educational Leadership, 62*(2), 90–91.

U.S. Department of Education, Office of Educational Research and Improvement. (1993). *What's wrong with writing and what can we do right now?* Washington, DC: Author.

Weiss, A. R., Lutkus, A. D., Grigg, W. S., & Niemi, R. G. (2001). The next generation of citizens: NAEP civics assessments: 1988 and 1998. *Education Statistics Quarterly, 3*(3), 21–24.

Weiss, A. R., Lutkus, A. D., Hildebrant, B. S., & Johnson, M. S. (2002). *The nation's report card: Geography 2001* (NCES 2002–484). Washington, DC: National Center for Education Statistics.

Wenglinsky, H.. (2000). *How teaching matters: Bringing the classroom back into discussions of teacher quality*. Princeton, NJ: Educational Testing Service, Policy Information Center.

Wenglinsky, H. (2002). The link between teacher classroom practices and student academic performance. *Education Policy Analysis Archives, 10*(2). Retrieved January 23, 2008, from http://epaa.asu.edu/epaa/v10n12/

Wenglinsky, H. (2003). Using large-scale research to gauge the impact of instructional practices on student reading comprehension: An exploratory study. *Education Policy Analysis Archives, 11*(9). Retrieved January 23, 2008, from http://epaa.asu.edu/epaa/v11n19/

Wenglinsky, H. (2004). The link between instructional practice and the racial gap in middle schools. *Research in Middle Level Education, 28*(1). Retrieved January 23, 2008, from http://www.nmsa.org/portals/0/pdf/publications/RMLE/rmle_vol28_no1_article1.pdf

NELS

Burkam, D. T., Lee, V. E., & Smerdon, B. A. (1997). Gender and science learning early in high school: Subject matter and laboratory experiences. *American Educational Research Journal, 34*(2), 297–331.

Butty, J.-A. L. M. (2001). Teacher instruction, student attitudes, and mathematics performance among 10th and 12th grade black and Hispanic students. *Journal of Negro Education, 70*(1/2), 19–37.

Carbonaro, W. J., & Gamoran, A. (2002). The production of achievement inequality in high school English. *American Educational Research Journal, 39*(4), 801–827.

Chaney, B. (1994). *Teachers' academic backgrounds and student outcomes in science and mathematics: NSF/NELS:88 teacher transcript analysis*. Rockville, MD: Westat.

Chaney, B. (1995). *Student outcomes and the professional preparation of eighth-grade teachers in science and mathematics: NSF/NELS:88 teacher transcript analysis.* Rockville, MD: Westat.

Gamoran, A., & Carbonaro, W. J. (2002/2003). High school English: A national portrait. *High School Journal, 86*(2), 1–13.

Gau, S.-J. (1997, April). *The distribution and the effects of opportunity to learn on mathematics achievement.* Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Hoffer, T. B., & Moore, W. M. (1996). *High school seniors' instructional experiences in science and mathematics: National Education Longitudinal Study of 1988*. Washington, DC: National Center for Education Statistics.

Horn, L., Hafner, A. L., & Owings, J. A. (1992). *National Education Longitudinal Study of 1988: A profile of American eighth-grade mathematics and science instruction*. Washington, DC: U.S. Government Printing Office.

Johnson, R. M. (2000, April). *Gender differences in mathematics performance: Walberg's educational productivity model and the NELS:88 database.* Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Lee, V. E., Smith, J. B., & Croninger, R. G. (1997). How high school organization influences the equitable distribution of learning in mathematics and science. *Sociology of Education, 70*(2), 128–150.

Rowan, B., Chiang, F. S., & Miller, R. J. (1997). Using research on employees' performance to study the effects of teachers on students' achievement. *Sociology of Education, 70*(4), 256–284.

Scott, L. A., & Ingels, S. J. (1992, April). *Transitions to high school: Instruction and achievement: Findings from the NELS:88 first follow-up (1990) student survey*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Smerdon, B. A., Burkam, D. T., & Lee, V. E. (1999). Access to constructivist and didactic teaching: Who gets it? Where is it practiced? *Teachers College Record, 101*(1), 5–34.

ECLS

Denton, K., West, J., & Walston, J. (2003). *Reading—Young children's achievement and classroom experiences: Findings from the Condition of Education 2003* (NCES 2003–070). Washington, DC: National Center for Education Statistics.

Judge, S., Puckett, K., & Cabuk, B. (2004). Digital equity: New findings from the Early Childhood Longitudinal Study. *Journal of Research on Technology in Education, 36*(4), 383–396.

Rathbun, A. H., & Hausken, E. G. (2003). Reading and writing instruction in kindergarten: How often and who receives it? Chicago: American Educational Research Association.

Rathbun, A. H., Walston, J. T., & Hausken, E. G. (2000, April). *Kindergarten teachers' use of developmentally appropriate practices: Results from the Early Childhood Longitudinal Study, kindergarten class of 1998–1999.* Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Rathbun, A. H., West, J., & Germino-Hausken, E. (2003). *Young children's access to computers in the home and at school in 1999 and 2000* (NCES 2003-36). Washington, DC: National Center for Education Statistics.

Walston, J., & West, J. (2004). *Full-day and half-day kindergarten in the United States: Findings from the Early Childhood Longitudinal Study, kindergarten class of 1998–99* (NCES 2004-078). Washington, DC: National Center for Education Statistics.

Zeng, G., & Zeng, L. (2005). Developmentally and culturally inappropriate practice in U.S. kindergarten programs: Prevalence, severity, and its relationship with teacher and administrator qualification. *Education, 125*(4), 706–724.