NATIONAL CENTER FOR EDUCATION EVALUATION AND REGIONAL ASSISTANCE
Institute of Education Sciences

INSTRUCTIONAL PRACTICES AND STUDENT MATH ACHIEVEMENT: CORRELATIONS FROM A STUDY OF MATH CURRICULA

This brief is directed to researchers and adds to the research base about instructional practices that are related to student achievement. Additional evidence on these relationships can suggest specific hypotheses for the future study of such instructional practices, which, in turn, will provide research evidence that could inform professional development of teachers and the writing of instructional materials. The results of this study revealed a pattern of relationships largely consistent with earlier research, but not in every case. Results that are consistent with previous research include increased student achievement associated with teachers dedicating more time to whole-class instruction, suggesting specific practices in response to students' work (1st grade only), using more representations of mathematical ideas, asking the class if it agrees with a student's answer, directing students to help one another understand mathematics, and differentiating curriculum for students above grade level (2nd grade only). Less consistent results were found in three 2nd-grade results, and include lower achievement associated with teachers' higher frequency of eliciting multiple strategies and solutions; prompting a student to lead the class in a routine; and with students more frequently asking each other questions. These findings suggest that practices associated with higher achievement gains include both student-centered and teacher-directed practices; however, some student-centered practices were associated with lower achievement gains.

How should teachers instruct young students in mathematics? Instructional practices have been proposed based on successful research projects (e.g., Clements and Sarama 2012; Montague and Jitendra 2012; National Mathematics Advisory Panel 2008; National Research Council 2009; Rohrer and Pashler 2010; Rosenshine 2012; National Association for the Education of Young Children 2013) and from observations of successful teachers (Lemov 2010). In addition to research suggesting practices that teachers should use, curricula by their nature also embody a set of instructional practices.

Prior studies have examined the relationships between various instructional practices and student achievement, but uncertainty remains about which practices teachers should use. For example, some studies compared practices that often are classified as "teacher-directed" with those often classified as "student-centered." Although some have argued that this is too simple a dichotomy (National Mathematics Advisory Panel 2008), we used it as a broad classificatory structure for two reasons. First and most important, the issue was prominent when the original study was designed. Indeed, it was the first issue addressed by the National Math Advisory Panel's (2008) instructional practices subpanel (for other examples of research that moves beyond the dichotomy, see Bodovski and Farkas 2007; Cohen and Hill 2000; Guarino et al. 2013; Le et al. 2006). Second, we believed our analyses could shed light on the nuances of the instructional practices included in those two broad categories, thus helping the field to move beyond the dichotomy if results so warranted.

The National Mathematics Advisory Panel (2008) defined teacher-directed instruction as "instruction in which primarily the teacher is communicating the mathematics to the students directly and in which the majority of interactions about the mathematics are between the teacher and the student" (pp. 6–14), and they defined student-centered instruction as "instruction in which primarily students are doing the teaching of the mathematics and that the majority of the interactions about the mathematics occurs between and among students" (pp. 6–16). Some of these studies found that teacher-directed practices led to greater student math achievement (Baker et al. 2002; Darch et al. 1984; Hopkins et al. 1997; Rittle-Johnson 2006), whereas other studies making similar comparisons found that student-centered practices were more effective (Bransford et al. 2000; Fuchs et al. 1995; Hickey et al. 2001; Muthukrishna and Borkowski 1995; Slavin and Karweit 1985) or had mixed results (Fuchs et al. 2006). Differences in findings and conclusions could be the result of important differences across the studies, including differences in the specific teacher-directed and student-centered practices examined.

Purpose of this Brief

This study builds on the research base by using correlational techniques to examine the relationships between a large number of specific instructional practices (39) and mathematics achievement of 1st- and 2nd-grade students. The instructional practices (described in the following chapter) are those used by teachers who participated in a randomized controlled trial (RCT) of four mathematics curricula: (1) Investigations in Number, Data, and Space (Investigations); (2) Math Expressions; (3) Saxon Math; and (4) Scott Foresman-Addison Wesley Mathematics (SFAW).[1] The four curricula use different approaches to student instruction, with differing levels of emphasis on various instructional practices (a detailed description of each curriculum appears in the Appendix). Those relative emphases may have guided study teachers' practice in ways that resulted in different student achievement. Alternatively, certain instructional practices may have been relatively more effective, regardless of curriculum.

The data from this study provide a unique opportunity to examine the relationship between individual instructional practices and student achievement. Some of the advantages of these data include independent observations of the same practices across all classrooms by observers not affiliated with any curricula approach; a highly reliable, individually administered student outcome assessment; a rich data set that includes controls for many characteristics of students, teachers, schools, and classroom environments; and a large sample, to examine whether the uses and relationships (to achievement) of the practices differ across 1st and 2nd grade. Characteristics of a grade may influence the way teachers use various practices. For example, students in 2nd grade may be able to work in groups, but such work may be more challenging for 1st graders (e.g., the cognitive load of dealing both with social interactions and the mathematics may be more likely to overwhelm the information-processing system of the younger students). These data are also unique, in that most of the schools and teachers were implementing their assigned curriculum for the first time, so prior experience with the recommended practices of each curriculum was comparable across teachers. Last, the random assignment of curricula to schools (and thus to the accompanying recommended practices within each curriculum) also may help guard against unmeasured variables related to teacher usage of particular practices.

The design of the study also has limitations relevant to interpreting the results. First and foremost, although the curricula were randomly assigned, the instructional practices observed were not. Therefore, this is a correlational study whose results must be evaluated with causal designs before implications for practice can be reliably drawn. Second, because there was limited variation in the use of some instructional practices, it is possible that a subset of these practices may have a true relationship with student achievement that could not be detected in these data. A third limitation is that the data were collected in single day-observations in each teacher's classroom.[2] Fourth, the observations recorded frequency but not the quality of implementation of the practices (see the discussion in Section II.A). Given these limitations, our results will be most useful in guiding hypotheses about relationships between specific instructional practices and student achievement.

In the next section, we present details about the results; in the following section, we interpret our results in the context of the broader literature to suggest why apparent inconsistencies have arisen. Such interpretations may help researchers design causal studies to accurately identify practices that teachers should or should not implement in classroom instruction—this study, as with most of the existing research, is based only on correlations. In the final section, we summarize our conclusions on what the current body of knowledge (including our findings) suggests about aspects of instruction that may promote students' learning of mathematics, and what research questions need to be addressed.

Relationships Between Instructional Practices and Student Math Achievement

Using data from the impact evaluation of four mathematics curricula, we examined the relationship between instructional practices and math achievement separately for students in 1st and 2nd grades. The 1st-grade sample includes 3,818 students from 364 classrooms across 108 schools; the 2nd-grade sample includes 2,796 students from 269 classrooms across 71 schools. Compared with all U.S. elementary schools, those included in the analyses by design have higher fractions of low-income and minority students and are more urban. The Appendix provides more information about the characteristics of schools, teachers, and students in the analysis.

Instructional Practices

Using a protocol developed by the study team, we measured a large number of instructional practices in each study classroom. The observations were conducted in the 2006–07 or 2007–08 schools years, during each school's first year of curriculum implementation in the study. The observations occurred within schools that were randomly assigned to use one of the four study curricula in both 1st and 2nd grades; about a quarter of the observations were conducted in each of the four study curricula.

The observation protocol includes nearly 100 items that were either thought to be useful by the study team for discriminating the instructional approaches of the study's four curricula, or were practices with prior evidence suggesting they are related to student achievement. The items were coded during one day's worth of math instruction in each study classroom that was observed in real time—this included the math lesson and the morning meeting or calendar time, which was typically about 70 minutes per day, on average. About two-thirds of the items on the protocol were coded during the observation; the remaining items were coded immediately after the observation.[3]

Observations of instructional practices were coded for frequency for two reasons. First and most important, low-inference observational measures are more reliable than high-inference measures (Brabeck et al. 2013, Campbell et al. 2004), and any introduction of Likert-based items rating quality would result in higher-inference items (as opposed to a simple count of the number of occurrences of each practice). Second, research in early mathematics education has not reliably identified the characteristics of low- and high-quality implementations of instructional practices (for a review, see Hiebert and Grouws 2007). Therefore, we recorded the frequency of specific practices to ascertain which are related to student achievement to guide future research that examines issues of quality and uses causal designs (which must identify a reasonable number of such practices to assign). More information about the development of the protocol is provided in the Appendix.

Groups of items on the protocol were designed to measure different concepts from slightly different perspectives; however, the study team's lack of prior experience using the protocol led the team to remain open to the possibility that the intended concepts may need to be revised once the observation data were examined. Thus, an exploratory factor analysis (EFA) was conducted by the study team to identify groupings. The EFA indicated that four factors underlie the observation data, and analyses indicated that scales developed for three of those factors were related to student achievement (Agodini et al. 2010).

The practices examined in this analysis are the 39 items that underlie the three scales the study team found are related to student achievement while working on the curriculum impact analysis.[4] One of the three scales measured practices that are typical in student-centered instructional settings, another measured practices typical in teacher-directed settings, and the last measured student interactions (peer collaboration) during math instruction. A fourth scale emerged from the EFA that included items about the classroom environment, such as the extent to which teachers need to manage student behavior. When calculating the relationship between each of the 39 practices and student achievement, we adjusted for the items in the classroom environment scale, as described below.[5] Agodini et al. (2010) provide more details about the process for constructing the scales.

While the terms "student-centered" and "teacher-directed" instruction have been used to describe quite different instructional practices and environments in the literature, it is important to remember that these terms are also used to describe the 39 items that clustered together in the EFA. Thus, the items in each scale may differ from those classified a priori in other studies, such as the report of the National Mathematics Advisory Panel (2008).

In Table 1, we list the 39 instructional practices examined, along with the various ways the items were coded. Most items asked observers to tally the frequency of the practice because the study's four curricula differ in the extent to which these practices are used. For these items, observers were instructed to stop tallying practices once the practice was observed 21 times. The cut-off was established for a variety of reasons, including limiting the burden placed on observers (see the Appendix for further information). The remaining items asked observers to indicate whether the practice did or did not occur (a yes/no coding) or the extent to which a statement (such as "students help one another understand math concepts or procedures") is characteristic of the class, because such coding of these items was considered sufficient for differentiating the study's curricula. The Appendix and Agodini et al. (2010) provide more details about the classroom observations.

Table 1. Instructional Practices in the Analysis

Item	Measurement	Possible Range
Items in the Student-Centered Instruction Scale		
Teacher poses open-ended questions that have more than one correct answer	Tally	0–21
Number of problems for which the teacher elicits multiple strategies or solutions	Tally	0–21
Teacher tells student the strategy to use in response to student work/answer	Tally	0–21
Teacher elicits other students' questions about a student's response	Tally	0–21
Teacher labels math strategy, problem, or concept in response to student work/answer	Tally	0–21
Teacher repeats student answer in a neutral way with no indication of correctness	Tally	0–21
Teacher probes for reasoning or justification in response to student work/answer	Tally	0–21
Teacher provides hint to students in response to student work/answer	Tally	0–21
Teacher clarifies what student says or does in response to student work/answer	Tally	0–21
Teacher extends what student says or does in response to student work/answer	Tally	0–21
Teacher uses praise or makes positive comments focused on content	Tally	0–21
Teacher highlights student work or solution to class	Tally	0–21
Number of different types of visual or three-dimensional representations created by students	Tally	0–21
Teacher differentiates curriculum for children who are above grade level	Scale	1–4
Items in the Teacher-Directed Instruction Scale		
Teacher asks close-ended questions	Tally	0–21
Number of problems for which the teacher guides practice on problems	Tally	0–21
Number of representations demonstrated by the teacher	Tally	0–21
Teacher indicates if correct without elaborating in response to student work/answer	Tally	0–21
Teacher calls on other students until the correct answer is given	Tally	0–21
Teacher asks class if it agrees or disagrees with a student's response	Tally	0–21
Teacher prompts student to guide practice or lead the class in a routine	Yes/No	0–1
Students practice number facts or procedures	Scale	0–6
Students provide choral or group responses to questions	Scale	0–2
Students rote count (orally or in writing)	Yes/No	0–1
Number of types of rote counting that occurred, by ones, by twos, and so forth	Check Box	0–8
Number of practice problems focusing on review of previously learned material	Tally	0–21
Number of materials used by children	Check Box	0–11
Number of types of representations used during math, by the teacher or by students	Check Box	0–7
Percentage of math instructional time spent in large group	Scale	0–4
Items in the Peer Collaboration Scale		
Teacher demonstrates how to play a game	Yes/No	0–1
Teacher directs or encourages students to help one another with math	Yes/No	0–1
Students play math games	Scale	0–6
Students ask peers questions about math	Scale	0–2
Students discuss math strategies or solutions with partner or small group	Scale	0–2
Percentage of math instructional time spent in small group	Scale	0–4
Percentage of math instructional time spent in pairs	Scale	0–4
Teacher encourages students to help one another understand math	Scale	1–4
Students help one another understand math concepts or procedures	Scale	1–4
Peer-to-peer interaction about math occurs	Scale	1–4

Because the coding scheme for the items supports frequency rather than quality of implementation of the practices (a limitation of the study), the practices may have been displayed across a range of implementation quality. In addition, for most items on the observation protocol, the frequency of observed practices is a function of the length of the lesson. Thus, our results will be most useful in guiding hypotheses about relationships between specific instructional practices and student achievement.

Student Math Achievement

Students' math achievement was measured by the study team administering the math assessment developed for the Early Childhood Longitudinal Study-Kindergarten Class of 1998–1999 (ECLS-K). The ECLS-K assessment is an adaptive and nationally normed test that meets accepted standards of validity and reliability (Rock and Pollack 2002). It also meets other important study requirements, including individual administration, ability to measure achievement gains over the study's grade range, and accuracy in capturing achievement of students from a wide range of backgrounds and ability levels. More details about the assessment, including reliability information for the study sample, appear in the Appendix.

Students were tested in both the fall and spring of the school year during which they participated in the study (2006–07 or 2007–08). The ECLS-K K–1 math assessment was administered to 1st graders. An ECLS-K math assessment for the 2nd grade did not exist; therefore, the study team worked with the Educational Testing Service (ETS), the developer of the ECLS-K, to create a 2nd-grade assessment by selecting appropriate items from existing ECLS-K math assessments (including the K–1, 3rd- , and 5th-grade instruments). ETS used information from the ECLS-K bridge study,[6] which included a small sample of 2nd graders, combined with information about the current study's sample, to ensure that the administered items appropriately targeted the estimated range of 2nd graders' ability levels.[7] The study also used a Spanish version of the assessment for any classes in which math instruction was conducted entirely in Spanish.

Analytic Approach

We used a two-step process to identify the instructional practices related to student achievement. The first step examined how each practice is related to student achievement when taking into account the statistical influence of other practices in the same scale. This involved three analyses—one for each set of items in the three scales that emerged from the EFA analyses of classroom observation data. For example, we examined the relationship between teacher use of open-ended questions (one of the 14 practices in the student-centered scale) and student achievement, adjusted for all other practices in the student-centered scale. Similarly, for the other two analyses, we examined the relationship between student achievement and the practices in the teacher-directed scale (15 items) and those in the peer collaboration scale (10 items). Each set of first-step results is adjusted for characteristics of schools, teachers, and students (including earlier achievement) and aspects of the classroom environment (such as the extent to which teachers managed student behavior).[8]

The second step involved an analysis that examined how practices identified in the first step relate to student achievement when such practices are considered simultaneously. Given that teachers frequently use practices from more than one instructional scale, the simultaneous consideration of those practices in each scale that correlate with student achievement may suggest those that matter in real-life situations. For example, suppose that the first step of the analysis shows that only one student-centered practice and one teacher-directed practice are significantly related to student achievement: teacher use of open- and close-ended questions. The second step would then examine the relationship between open-ended questions and student achievement adjusted for close-ended questions, and vice versa. As with the first-step results, the second-step results are adjusted for characteristics of schools, teachers, and students and aspects of the classroom environment.[9]

We conducted all analyses separately for 1st- and 2nd-grade students for two main reasons. First, we did so because the relationship of the instructional practices could depend on student maturity (e.g., the capacity to attend both to content and social-cognitive demands of productive cooperative learning change substantially over the early years (Nastasi and Clements 1991)). Second, the mathematical focus of the grade level is different, with new emphases on mathematical structures, such as units in place value and length measure, emerging in 2nd grade. The Appendix includes more details about the analytic approach.

We used this analytic approach because practices that are consistent with a particular instructional approach, such as teacher-directed practices, tend to co-occur. But, within a set of practices, there may be ones that are and are not related to student achievement. Including in each of the first-step analyses all the practices that are consistent with a particular approach may help identify the ones in each domain that have the greatest correlation with achievement. The second step then examines which of the identified practices in the various domains have the greatest correlation with achievement.

Limitations of the Analysis

It is important to note three limitations of our analysis. First, the analytic approach does not support causal statements about the achievement effect of the examined practices. Teachers who used certain instructional practices may differ in ways that affect achievement but are not contained in our data source; therefore, we could not account for these differences when examining the relationship between the practices and student achievement. For example, a practice associated with an increase in student achievement could reflect not only the effect of the practice but also the possibility that more effective teachers adopted the practice—and we did not measure teacher effectiveness. However, as Hiebert and Grouws (2007, p. 397) stated, correlation approaches "can help to identify features of teaching that might have major effects on learning and deserve further study, and they can provide descriptive data to inform the development of hypotheses regarding teacher-learning relationships."

Second, there is limited variation in the use of some instructional practices, so it is possible that a subset of these practices may have a true relationship with student achievement that could not be detected in these data. One reason for the limited variation could be due to the way in which some items were coded. For example, as mentioned, items tallied by observers were top coded to 21 if teachers exhibited those practices at least 21 times, thus introducing a ceiling on the tallies. As shown in Table 2, about 95 percent of teachers received the maximum code for the item that tallied the number of times a teacher asked a close-ended question. Our finding in this brief (described below) that this teacher practice is not related to student achievement could be due to the limited variation in the practice brought on by our approach for coding the item.

Another reason for the limited variation for some items could be due to the influence of the curricula included in the RCT on teachers' practices. As mentioned in Agodini et al. (2010), nearly all teachers reported use of the curriculum assigned to their schools. Therefore, practices that are not part, or are a minor part, of the study's curricula may have a low frequency of use. A final reason for the limited variation for some items is that most teachers were using the curriculum in the RCT for the first time when classroom observations were conducted. During the first year of curriculum implementation, teachers may be learning to implement the structure and procedures of the curriculum and may not implement all parts (or the more sophisticated parts) of the curriculum until they gain experience and reach a more refined phase of implementation. Therefore, more challenging practices may have a low frequency of use.

Third, the data were collected in single-day observations in each teacher's classroom.[10] All 1st-grade observations occurred in the spring of the school year, and the 2nd-grade observations were conducted in fall, winter, and spring. However, as single-day observations, they do not necessarily reflect teacher practice across the entire school year. In addition, the observations were conducted in real time, so observers could not "replay" an interaction to see if a practice occurred. This could result in measurement error, which could weaken the correlations. Given these limitations, the results presented in the following section are best interpreted as suggesting hypotheses for further study.

Results

In Tables 3 through 5, we present results from the first step and, in Table 6, results from the second step; we focus on results from the second step. In both steps, practices that are statistically significant at the 0.10 level are deemed to be associated with student achievement. We use the 0.10 level of statistical significance rather than 0.05 to avoid false conclusions about practices that may be related to student achievement but do not reach the conventional level because of little variation in our data.[11] Considering the correlational design of this study and the exploratory nature of the results, we suggest that each of our findings be explored with causal designs. Given that researchers will most likely study only those practices that are associated with achievement, we believe it is important to minimize Type II error (failing to detect associations), even though this may lead to increased Type I error (detecting associations that really do not exist), so more rigorous future research does not overlook practices that are potentially important for student achievement.

Table 2. Instructional Practices

Item	Average	Minimum	Percentile 25th	50th	75th	Maximum
Items in the Student-Centered Instruction Scale						
Teacher poses open-ended questions that have more than one correct answer (TALLY)	6.18	0	1	4	9	21
Number of problems for which the teacher elicits multiple strategies or solutions (TALLY)	1.83	0	0	1	3	11
Teacher tells student the strategy to use in response to student work/answer (TALLY)	1.39	0	0	1	2	21
Teacher elicits other students' questions about a student's response (TALLY)	0.18	0	0	0	0	6
Teacher labels math strategy, problem, or concept in response to student work/answer (TALLY)	1.39	0	0	0	2	21
Teacher repeats student answer in a neutral way with no indication of correctness (TALLY)	1.56	0	0	0	2	21
Teacher probes for reasoning or justification in response to student work/answer (TALLY)	4.71	0	1	4	7	21
Teacher provides hint to students in response to student work/answer (TALLY)	6.74	0	2	5	9	21
Teacher clarifies what student says or does in response to student work/answer (TALLY)	1.69	0	0	1	2	21
Teacher extends what student says or does in response to student work/answer (TALLY)	0.63	0	0	0	1	21
Teacher uses praise or makes positive comments focused on content (TALLY)	1.96	0	0	1	3	21
Teacher highlights student work or solution to class (TALLY)	0.98	0	0	0	1	21
Number of different types of visual or three-dimensional representations created by students (TALLY)	2.13	0	1	2	3	15
Teacher differentiates curriculum for children who are above level (SCALE, 1–4)	1.16	1	1	1	1	4
Items in the Teacher-Directed Instruction Scale						
Teacher asks close-ended questions (TALLY)	20.14	1	21	21	21	21
Number of problems on which the teacher guides practice on problems (TALLY)	9.19	0	4	8	14	21
Number of representations demonstrated by the teacher (TALLY)	7.38	0	3	6	10	21
Teacher indicates if correct without elaborating in response to student work/answer (TALLY)	18.93	0	21	21	21	21
Teacher calls on other students until the correct answer is given (TALLY)	2.67	0	0	2	4	21
Teacher asks class if it agrees or disagrees with a student's response (TALLY)	2.19	0	0	1	3	21
Teacher prompts student to guide practice or lead the class in a routine (YES/NO)	0.33	0	0	0	1	1
Students practice number facts or procedures (SCALE, 1–6)	3.37	0	0	6	6	6
Students provide choral or group responses to questions (SCALE, 0–2)	1.27	0	1	1	2	2
Students rote count (orally or in writing) (YES/NO)	0.63	0	0	1	1	1
Number of types of rote counting that occurred, by ones, by twos, and so forth (TOTAL OF 8 TIMES)	1.78	0	0	2	3	7
Number of practice problems focusing on review of previously learned material (TALLY)	7.01	0	0	4	11	21
Number of materials used by children (TOTAL OF 11 ITEMS)	1.67	0	1	2	2	6
Number of types of representations used during math, by the teacher or by students (TOTAL OF 7 ITEMS)	2.26	0	1	2	3	7
Percentage of math instructional time spent in large group (SCALE, 0–4)	3.11	0	3	3	4	4

Table 2 *(continued)*

Item	Average	Minimum	Percentile 25th	Percentile 50th	Percentile 75th	Maximum
Items in the Peer Collaboration Scale						
Teacher demonstrates how to play a game (YES/NO)	0.15	0	0	0	0	1
Teacher directs or encourages students to help one another with math (YES/NO)	0.41	0	0	0	1	1
Students play math games (SCALE, 0–6)	1.18	0	0	0	0	6
Students ask peers questions about math (SCALE, 0–2)	0.39	0	0	0	1	2
Students discuss math strategies or solutions with partner or small group (SCALE, 0–2)	0.48	0	0	0	1	2
Percentage of math instructional time spent in small group (SCALE, 0–4)	0.25	0	0	0	0	4
Percentage of math instructional time spent in pairs (SCALE, 0–4)	0.51	0	0	0	1	4
Teacher encourages students to help one another understand math (SCALE, 1–4)	1.72	1	1	1	2	4
Students help one another understand math concepts or procedures (SCALE, 1–4)	1.78	1	1	2	2	4
Peer-to-peer interaction about math occurs (SCALE, 1–4)	1.71	1	1	2	2	4

Source: Author calculations (means and quartile ranges) using classroom observation data.

Note: TALLY has a possible range of 0–21, where 21 includes tallies of 21 and over.

Table 3. Practices in the Student-Centered Scale: Percentile-Point Increase in Student Achievement Associated with a One-Unit Increase in Each Practice (Step One Results)

	1st Grade		2nd Grade	
Item	Estimate	*p*-value	Estimate	*p*-value
Teacher poses open-ended questions that have more than one correct answer (TALLY)	-0.1	0.52	0.2*	0.06
Number of problems for which the teacher elicits multiple strategies or solutions (TALLY)	0.0	0.95	-0.8**	0.04
Teacher tells student the strategy to use in response to student work/answer (TALLY)	0.8**	0.03	-0.1	0.66
Teacher elicits other students' questions about a student's response (TALLY)	-0.8	0.49	0.8	0.33
Teacher labels math strategy, problem, or concept in response to student work/answer (TALLY)	0.5	0.12	0.2	0.42
Teacher repeats student answer in a neutral way with no indication of correctness (TALLY)	-0.3	0.19	-0.4*	0.08
Teacher probes for reasoning or justification in response to student work/answer (TALLY)	0.0	0.87	0.1	0.62
Teacher provides hint to students in response to student work/answer (TALLY)	-0.2*	0.10	0.1	0.23
Teacher clarifies what student says or does in response to student work/answer (TALLY)	0.2	0.61	-0.1	0.69
Teacher extends what student says or does in response to student work/answer (TALLY)	0.2	0.67	-0.1	0.86
Teacher uses praise or makes positive comments focused on content (TALLY)	0.0	0.88	0.2	0.48
Teacher highlights student work or solution to class (TALLY)	0.2	0.45	0.1	0.71
Number of different types of visual or three-dimensional representations created by students (TALLY)	0.3	0.44	0.2	0.57
Teacher differentiates curriculum for children who are above grade level (SCALE, 1–4)	-1.3	0.33	1.8*	0.09

Source: Author calculations using data from the fall and spring ECLS-K math test administered by the study team, classroom observations conducted by the study team, teacher surveys data collected by the study team, and Common Core Data.

Note: TALLY has a possible range of 0–21, where 21 includes tallies of 21 and over. The estimates were produced using a three-level hierarchical learning model (HLM); a list of covariates is provided in the Appendix.

*Significantly different from zero at the 0.10 level, two-tailed test.

**Significantly different from zero at the 0.05 level, two-tailed test.

***Significantly different from zero at the 0.01 level, two-tailed test.

Table 4. Practices in the Teacher-Directed Scale: Percentile Point Increase in Student Achievement Associated with a One-Unit Increase in Each Practice (Step One Results)

	1st Grade		2nd Grade	
Item	Estimate	*p*-value	Estimate	*p*-value
Teacher asks close-ended questions (TALLY)	-0.1	0.76	0.0	0.86
Number of problems on which the teacher guides practice on problems (TALLY)	0.0	0.84	0.0	0.91
Number of representations demonstrated by the teacher (TALLY)	0.0	0.92	0.3*	0.08
Teacher indicates if correct without elaborating in response to student work/answer (TALLY)	0.1	0.79	0.2	0.34
Teacher calls on other students until the correct answer is given (TALLY)	-0.3*	0.09	0.0	0.90
Teacher asks class if it agrees or disagrees with a student's response (TALLY)	0.1	0.54	0.3*	0.08
Teacher prompts student to guide practice or lead the class in a routine (YES/NO)	-1.2	0.39	-3.0*	0.05
Students practice number facts or procedures (SCALE, 1–6)	-0.2	0.43	0.2	0.44
Students provide choral or group responses to questions (SCALE, 0–2)	-0.9	0.39	-1.4	0.16
Students rote count (orally or in writing) (YES/NO)	-0.2	0.91	-1.3	0.47
Number of types of rote counting that occurred, by ones, by twos, and so forth (TOTAL OF 8 TIMES)	0.7	0.30	0.7	0.36
Number of practice problems focusing on review of previously learned material (TALLY)	0.2	0.11	0.0	0.76
Number of materials used by children (TOTAL OF 11 ITEMS)	-1.0	0.13	0.1	0.92
Number of types of representations used during math, by teacher or students (TOTAL OF 7 ITEMS)	0.0	0.95	0.1	0.86
Percentage of math instructional time spent in large group (SCALE, 0–4)	1.5**	0.04	-0.4	0.54

Source: Author calculations using data from the fall and spring ECLS-K math test administered by the study team, classroom observations conducted by the study team, teacher surveys data collected by the study team, and Common Core Data.

Note: TALLY has a possible range of 0–21, where 21 includes tallies of 21 and over. The estimates were produced using a three-level HLM; a list of covariates is provided in the Appendix.

*Significantly different from zero at the 0.10 level, two-tailed test.

**Significantly different from zero at the 0.05 level, two-tailed test.

***Significantly different from zero at the 0.01 level, two-tailed test.

Table 5. Practices in the Peer Collaboration Scale: Percentile Point Increase in Student Achievement Associated with a One-Unit Increase in Each Practice (Step One Results)

	1st Grade		2nd Grade	
Item	Estimate	*p*-value	Estimate	*p*-value
Teacher demonstrates how to play a game (YES/NO)	1.6	0.56	4.3	0.14
Teacher directs or encourages students to help one another with math (YES/NO)	1.2	0.43	1.0	0.54
Students play math games (SCALE, 0–6)	-0.5	0.26	-0.3	0.48
Students ask peers questions about math (SCALE, 0–2)	0.0	0.99	-4.2***	0.01
Students discuss math strategies or solutions with partner or small group (SCALE, 0–2)	1.7	0.25	-0.9	0.52
Percentage of math instructional time spent in small group (SCALE, 0–4)	-0.6	0.56	-0.2	0.83
Percentage of math instructional time spent in pairs (SCALE, 0–4)	0.8	0.37	0.6	0.53
Teacher encourages students to help one another understand math (SCALE, 1–4)	0.3	0.78	1.2	0.36
Students help one another understand math concepts or procedures (SCALE, 1–4)	-2.4*	0.10	2.3*	0.08
Peer-to-peer interaction about math occurs (SCALE, 1–4)	-0.9	0.50	-1.6	0.28

Source: Author calculations using data from the fall and spring ECLS-K math test administered by the study team, classroom observations conducted by the study team, teacher surveys data collected by the study team, and Common Core Data.

Note: The estimates were produced using a three-level HLM; a list of covariates is provided in the Appendix.

*Significantly different from zero at the 0.10 level, two-tailed test.

**Significantly different from zero at the 0.05 level, two-tailed test.

***Significantly different from zero at the 0.01 level, two-tailed test.

In the 1st grade, greater use of the following practices is associated with an increase in math achievement (Table 6):

- Teachers telling students the strategy to use in response to students' work or answers
- Higher percentage of math instructional time spent in a large-group or whole-class setting

In the 2nd grade, greater use of the following practices is associated with an increase in math achievement:

- Teachers differentiating curriculum for children who are above grade level.
- Number of representations that teachers demonstrate
- Teachers asking the class if it agrees with a student's answer
- Students help one another understand math concepts or procedures

Unlike the case of the 1st-grade results, however, some practices in 2nd grade are associated with a *decrease* in math achievement, including:

- Teachers eliciting multiple strategies or solutions
- Teachers prompting a student to guide practice or lead the class in a routine
- Frequency of students asking one another questions about math

A description of how each of these items was coded is provided in the Appendix.

Placing Our Results in Context With the Findings of Similar Studies

A relatively large body of research has attempted to identify instructional practices related to student achievement (e.g., Hiebert and Grouws 2007). However, this research has faced many difficulties because many factors, both inside and outside the classroom, are likely to affect student achievement. Much of the earlier research is based on correlational methods similar to those used in the current study, and those methods may not adequately account for all the factors affecting student achievement. In addition, even though some of the earlier research used rigorous methods (such as random assignment to specific practices), such research is sparse and can involve small sample sizes. This study has the advantages of a large and rich data set, independent observations of the same practices across classrooms randomly assigned to different curricula, and a reliable and valid student outcome assessment. However, because the practices themselves were not randomly assigned, the usefulness of the findings lies mainly in suggesting hypotheses for further study.

Table 6. Step Two Results: Percentile Point Increase in Student Achievement Associated with a One-Unit Increase in the Practices Identified in Step One (Tables 3, 4, and 5)

Item	Estimate	*p*-value
1st Grade		
Teacher tells student the strategy to use in response to student work/answer (TALLY)	0.9**	0.01
Teacher provides hint to students in response to student work/answer (TALLY)	-0.2	0.18
Teacher calls on other students until the correct answer is given (TALLY)	-0.3	0.16
Percentage of math instructional time spent in large group (SCALE, 0–4)	1.4**	0.04
Students help one another understand math concepts or procedures (SCALE, 1–4)	-1.0	0.24
2nd Grade		
Teacher poses open-ended questions that have more than one correct answer (TALLY)	0.2	0.14
Number of problems for which the teacher elicits multiple strategies or solutions (TALLY)	-0.7*	0.08
Teacher repeats student answer in a neutral way with no indication of correctness (TALLY)	-0.3	0.18
Teacher differentiates curriculum for children who are above grade level (SCALE, 1–4)	3.0***	0.01
Number of representations demonstrated by the teacher (TALLY)	0.3**	0.02
Teacher asks class if it agrees or disagrees with a student's response (TALLY)	0.3*	0.05
Teacher prompts student to guide practice or lead the class in a routine (YES/NO)	-2.7*	0.07
Students ask peers questions about math (SCALE, 0–2)	-4.1***	0.00
Students help one another understand math concepts or procedures (SCALE, 1–4)	2.1**	0.02

Source: Author calculations using data from the fall and spring ECLS-K math test administered by the study team, classroom observations conducted by the study team, teacher surveys data collected by the study team, and Common Core Data.

Note: TALLY has a possible range of 0–21, where 21 includes tallies of 21 and over. The estimates were produced using a three-level HLM; a list of covariates is provided in the Appendix.

*Significantly different from zero at the 0.10 level, two-tailed test.

**Significantly different from zero at the 0.05 level, two-tailed test.

***Significantly different from zero at the 0.01 level, two-tailed test.

To serve that purpose as well as possible, we place our results in the context of the findings of other studies that focused on the relationships between instructional practices and student achievement. In this way, we hope to enhance this research base and suggest specific hypotheses, with the goal of helping researchers design studies that will identify practices teachers should implement or avoid in their classroom instruction. Therefore, we examine what the body of knowledge (including our findings) suggests as hypotheses to pursue in establishing relationships between particular practices and student achievement.

As the discussion below explains, some of our findings are consistent with earlier research, whereas other findings are inconsistent. When possible, we note whether our findings are consistent or inconsistent with research that draws on more rigorous methods. In addition, we highlight instructional practices for which the research base seems to be nearing a consensus and discuss what the findings may mean for instruction, professional development, and curriculum development and refinement, if validated by experimental means.

Findings Consistent with Earlier Research

Move beyond a simple student-centered/teacher-directed dichotomy. One main finding that is consistent with earlier research is that the categorization into student-centered versus teacher-directed may not be a dichotomy on which researchers should base future studies. That is, our results are congruous with the conclusion of the National Mathematics Advisory Panel: "High-quality research does not support the contention that instruction should be either entirely 'student-centered' or 'teacher-directed.' Research indicates that some forms of particular instructional practices can have a positive impact under specified conditions" (2008, p. 11). Therefore, we will not make recommendations for studies that compare those broad categories of approaches, but rather will focus on hypotheses regarding which student-centered and teacher-directed practices appear to either warrant experimental validation (because they appear promising from the present research base) or need further research to resolve conflicting evidentiary claims or alternative interpretations.

Large group or whole-class instruction is useful. The finding that an increase in the percentage of math instructional time spent in a large group or whole-class setting is associated with an increase in math achievement has at least two possible interpretations. First, it may reflect the positive effects of active teaching, which features the teacher's presentation of what is to be learned and provision of individual feedback (Brophy 1988; Good et al. 1983). This finding would be consistent with the finding of a negative relationship between math achievement and the frequency with which teachers prompt a *student* to lead the class. Second, the positive relationship between a large group or whole-class setting and math achievement may also simply indicate that students in classes with more whole-class time devoted to math experience more instructional time in math ("time on task"; Bodovski and Farkas 2007; Brophy 1988; Pianta et al. 2008; Sylva et al. 2005; Wenglinsky 2004).

Future research could investigate the two potential interpretations and examine when other pedagogical structures (i.e., size and nature of student groups and the instructional practices used in each grouping) are, or are not, useful. For example, *guided* small groups—those in which a teacher is present, leading, monitoring, and guiding children's work—may complement whole-group instruction, but only if the tasks assigned to students guide them to productive engagement with important mathematical concepts and procedures (Hiebert and Grouws 2007).

Representations are useful if used appropriately. The study's results also address teachers' practices during instructional time and how those practices relate to student achievement. For example, we find that the more 2nd-grade teachers demonstrated representations of mathematical ideas, the greater the increase in student achievement. The result suggests the hypothesis that curricula and professional development will be more effective if they emphasize that aspect of "pedagogical content knowledge":

> [A] "particular form of content knowledge that embodies the aspects of content most germane to its teachability. . .[including] the most useful forms of representation of those ideas, the most powerful analogies, illustrations, examples, explanations, and demonstrations—in a word, the ways of representing and formulating the subject that make it comprehensible to others. Since there are no single most powerful forms of representation, the teacher must have at hand a veritable armamentarium of alternative forms of representation, some of which derive from research whereas others originate in the wisdom of practice" (Shulman 1986, p. 9).

However, a specific caveat is in order. The present study's data on representations simply measures the number of representations, not the quality or appropriateness of the representations. Teachers may have used some representations that were powerful and others that were not. The use of more representations may just have increased the likelihood that some of them were powerful. It is possible that targeted use of a smaller number of effective representations may be as beneficial as the use of a larger number of representations. Future research could first evaluate the effectiveness of different representations and then address the effects of different numbers of those shown to be efficacious. Research could also help identify the most effective representations.

It may be useful to structure students' strategies for particular learning objectives. A related issue involves the strategies children use and are taught to use. Most 1st graders and many 2nd graders are forming a collection of mathematical strategies for solving problems and figuring out arithmetic combinations (Sarama and Clements 2009). Our results suggest the hypothesis that it might be helpful for teachers to suggest specific strategies in response to students' work (such as has been shown to be successful in Japan and the U.S. (Henry and Brown 2008; Murata and Fuson 2006).

Students interacting and thinking about one another's work may be important. Also consistent with other research, our results demonstrate a positive relationship between teachers asking the class if it agrees with a student's answer and student achievement. Several researchers (Clements and Sarama 2009; Kamii and Dominick 1997) and position statements, such as the Curriculum Focal Points (National Council of Teachers of Mathematics 2006) and the Common Core State Standards (CCSSO/NGA 2010), suggest that students should "develop, discuss, and use" strategies to solve problems (e.g., Baroody and Rosu 2004). A careful analysis of survey data concluded that having 1st graders work problems and explain their problem solving is associated with (and probably causes) increased mathematics achievement in 1st grade (Guarino et al. 2013). Teachers promote discussion among students about these strategies by asking them to agree or disagree with a given student's answer (Kamii and Dominick 1997), and provide mathematical structure and language when appropriate. Alternatively, the agreement with a student's answer could be a method for more actively involving or focusing students when it is not their turn to solve a problem in large group. Future research will need to ascertain if such discussion—beyond simple agreement with an answer—is a factor in the association found in the present study and whether specific interventions can engender such facilitative discussion.

Differentiating instruction increases achievement of higher-performing students. Another finding consistent with earlier research demonstrates that an increase in differentiating curriculum for students who are above grade level is associated with an increase in student achievement. This is consistent with the finding of the National Mathematics Advisory Panel (2008), which reviewed studies on differentiating instruction for students who are gifted in mathematics and concluded that such differentiation benefits these students. The practice in the current analysis was characterized as "Teacher differentiates curriculum for children who are above grade level." Based on the results of this analysis, differentiation is a practice that predicts achievement (statistically significant in 2nd grade); however, it is a low-occurrence practice in this study's data. Accordingly, differentiation may be a particularly important practice if it relates to achievement gains, even when used infrequently. Moreover, the low rate of differentiation is consistent with the literature on higher-competence students and thus has implications for professional development. For example, in one study, observations of early childhood teachers show that such teachers usually misjudge children's level of mathematical thinking and therefore give practice ("more of the same") problems even when they intend to provide learning opportunities (challenging problems), especially to the highest-performing children (Bennett et al. 1984).

Findings Inconsistent with Earlier Research

A few of our findings are inconsistent with earlier research. For example, our results indicate a negative relationship between student achievement and the frequency with which teachers elicited the use of multiple strategies and solutions. Research previously cited shows, in contrast, that children benefit from inventing their own strategies and discussing a range of strategies for solving particularly demanding arithmetic problems. For example, the number of strategies discussed and applied by children predicts their later learning (Siegler 1995). As another example, when learning to solve problems involving mathematical equivalence, children were most successful when they had passed through a stage involving their consideration of several solution strategies (Alibali et al. 1993; Siegler 1995). Thus, our finding may be an anomaly resulting from the low frequency of this behavior. Or, children who experience the practice with limited frequency may not reach a skill level that allows them to consider different strategies and solutions. It may also be possible that encouraging children to use several strategies confuses students or wastes time unless the teacher successfully elicits articulation and defense of strategies, extends students' descriptions of those strategies, expects students to provide reasons and justifications for why those strategies are mathematically useful and valid, and provides mathematical structure and language as necessary (Fraivillig et al. 1999). Teachers in this study may have been prompted by the curriculum assigned to them but may not have yet developed such a comprehensive set of elicitation and support practices. Additional research is needed to disentangle these possibilities.

A related matter is the issue of students' interactions with one another; again, our results appear contradictory to several earlier studies. We did find, however, that one measure of student interaction—students helping one another understand math concepts or procedures—is positively associated with student achievement and consistent with previous research, which indicates that students benefit from interacting with peers regarding mathematics, mathematical problems, and their strategies and solutions, especially when the teacher builds expectations that students should help one another understand math concepts or procedures (Cobb and McClain 2002; Lampert and Cobb 2003; Stein et al. 2008; Johnson and Johnson 2009). Nonetheless, several other items we examined regarding peer collaboration are not associated with increased achievement. Indeed, "students more frequently asking each other questions" is associated with a decrease in student achievement. Generating multiple strategies and talking with peers takes time, and the amount of time spent on math instruction in the study classrooms may not have been sufficient for these practices to occur to a meaningful degree. Or, it may be that asking questions without getting answers is frustrating and unhelpful; that is, it may be that the *explanations* might make a difference (Fuchs et al. 1997; Nastasi and Clements 1991; Webb 1984). Again, future research is needed to address these alternative hypotheses, and we recommend that no implications be drawn for practice until such issues are addressed. At this point, our results suggest that collaborative learning practices should follow research-based recommendations designed to ensure that collaborative time is productive (Fuchs et al. 2001; Greenwood et al. 1989; Johnson and Johnson 2009; Nastasi and Clements 1991).

Conclusions

Our results, together with the results of other studies, suggest there are aspects of mathematics instructional environments, including teacher, student, and peer-to-peer components, that are statistically related to student achievement. Specifically, 1st-grade student achievement was higher when 1st-grade teachers increased the time devoted to whole-class instruction and in classrooms in which teachers suggested specific strategies in response to students' work. Second-grade student achievement was higher when 2nd-grade teachers used more representations of mathematical ideas, asked the class whether it agreed with a specific answer, encouraged students to help one another understand mathematics, and when teachers differentiated curriculum for students who were above grade level.

Because individual instructional practices do not operate in isolation, and because teachers tend to use more than one practice, these findings suggest that it may be useful for future research to focus on the effects of combinations of instructional practices, including those viewed as aligned with different pedagogical styles (such as student-centered and teacher-directed). Instructional practices need not function as opposing pedagogical styles or practices (as the National Mathematics Advisory Panel originally posed the question, and as many observational measures are structured), but rather as synergistic components. Therefore, future research ought to progress beyond dichotomies such as "student-centered versus teacher-directed instruction" and instead investigate how a variety of approaches from both perspectives could play complementary functions, and at which grade levels and with what content each is more effective. As one example, a study of peer-mediated instruction (PMI) showed strong positive effects compared to teacher-directed instruction (Fuchs et al. 1997). However, when teachers provided instruction on methods of presenting conceptual mathematical explanations within PMI, the effects were even stronger. Thus, the two approaches worked synergistically to engender the highest achievement. Similarly, our results for 2nd grade suggest that students can profitably help each other with math concepts or procedures, but it may be beneficial if the teacher guides the curriculum (e.g., representations used and differentiation) and classroom routines.

Because teachers were randomly assigned to implement one of the study's four curricula, which draw on the various instructional practices in different ways, and because implementation analyses presented in other study reports show that teachers tended to adhere to their assigned approach (Agodini et al. 2009, 2010), these study data do not provide a good opportunity to examine combinations of instructional practices outside of the four curricula. Therefore, we believe it would be more useful for future research to examine this issue.

It also would be productive to investigate the effects of other features of classroom environments and teaching that are not classified by the above dichotomies, such as the extent to which teachers and students attend explicitly to concepts and the extent to which students are challenged to engage in and struggle with important mathematics concepts (Clements and Sarama 2012, Hiebert and Grouws 2007).

Finally, in addition to looking at the effects of combined instructional practices, researchers could examine the difficulty of implementing various instructional practices. Recent work by Remillard (2012) focuses on the relative ease of implementing various curricula that embed combinations of practices. It could be instructive to further this line of research, along with research on the relative ease of implementing specific instructional practices. For example, our findings suggest that increasing the amount of math instructional time in a large group or whole-class setting is related to increased student achievement. Making this instructional change could be easy for many teachers. However, increasing the frequency with which students help one another understand math concepts or procedures—another practice positively associated with student achievement—could be more challenging for some teachers (particularly in some classrooms), in that it might require a higher level of pedagogical content knowledge than simply restructuring the class for more whole-class activities. Experiments on these issues would make a substantive contribution to both research and practice.

REFERENCES

Agodini, Roberto, Barbara Harris, Neil Seftor, Janine Remillard, and Melissa Thomas. "After Two Years, Three Elementary Math Curricula Outperform a Fourth." NCEE 2013-4019. Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education, 2013.

Agodini, Roberto, Barbara Harris, Melissa Thomas, Robert Murphy, and Lawrence Gallagher. "Achievement Effects of Four Early Elementary School Math Curricula: Findings for First and Second Graders." NCEE 2011-4001. Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education, 2010.

Agodini, Roberto, Barbara Harris, Melissa Thomas, Robert Murphy, and Lawrence Gallagher. "Achievement Effects of Four Early Elementary School Math Curricula: Findings for First Graders in 39 Schools." NCEE 2009-4052. Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education, 2009.

Agodini, Roberto, John Deke, Sally Atkins-Burnett, Barbara Harris, and Robert Murphy. "Design for the Evaluation of Early Elementary School Mathematics Curricula." Princeton, NJ: Mathematica Policy Research, January 2008.

Alibali, M.W., and S. Goldin-Meadow. "Gesture-Speech Mismatch and Mechanisms of Learning: What the Hands Reveal About a Child's State of Mind." *Cognitive Psychology,* vol. 25, no. 4, October 1993, pp. 468–573.

Baker, S., R. Gersten, and D-S Lee. "A Synthesis of Empirical Research on Teaching Mathematics to Low-Achieving Students." *Elementary School Journal,* vol. 103, no. 1, September 2002, pp. 51–73.

Baron, R.M., and D. Kenney. "The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations." *Journal of Personality and Social Psychology*, vol. 51, no. 6, 1986, pp. 1173–1182.

Baroody, A.J., and L. Rosu. "Adaptive Expertise with Basic Addition and Subtraction Combinations—the Number Sense View." Presented at the American Educational Research Association, San Francisco, CA, April 2004.

Bennett, N., C. Desforges, A. Cockburn, and B. Wilkinson. *The Quality of Pupil Learning Experiences*. Hillsdale, NJ: Erlbaum, 1984.

Bodovski, K., and G. Farkas. "Mathematics Growth in Early Elementary School: The Roles of Beginning Knowledge, Student Engagement, and Instruction." *The Elementary School Journal*, vol. 108, no. 2, November 2007, pp. 115–130.

Borman, G.D. "Designing Field Trials of Educational Interventions." In *Scale-Up in Education, Volume II: Issues in Practice*, edited by B. Schneider and S.-K. McDonald. Lanham, MD: Rowman & Littlefield, 2007.

Brabeck, M., C.A. Dwyer, K. Geisinger, R. Marx, G. Noell, R. Pianta, and F.C. Worrell. "Applying Psychological Science to Using Data for Continuous Teacher Preparation Program Improvement." Draft Report of the Board of Educational Affairs Task Force. Washington, DC: Board of Educational Affairs, American Psychological Association, 2013.

Bransford, J.D., A.L. Brown, and R.R. Cocking (eds.). *How People Learn*. Washington, DC: National Academy Press, 2000.

Brophy, J. "Research on Teacher Effects: Uses and Abuses." *The Elementary School Journal,* vol. 89, no. 1, September 1988, pp. 3–21.

Campbell, J., L. Kyriakides, D. Muijs, and W. Robinson. *Assessing Teacher Effectiveness: Developing a Differentiated Model.* New York, NY: RoutledgeFalmer, 2004.

CCSSO/NGA. "Common Core State Standards for Mathematics." Washington, DC: Council of Chief State School Officers and the National Governors Association Center for Best Practices. 2010. Available at [http://corestandards.org/]. Accessed August 30, 2013.

Charles, Randall, Janet H. Caldwell, Mary Cavanagh, Dinah Chancellor, Juanita V. Copley, Warren D. Crown, Francis (Skip) Fennell, Alma B. Ramirez, Kay B. Sammons, Jane F. Schielack, William Tate, and John A. Van de Walle. *Scott Foresman-Addison Wesley enVision Math.* Glenview, IL: Pearson Scott Foresman, 2009.

Charles, Randall, Warren Crown, Francis (Skip) Fennell, Janet H. Caldwell, Mary Cavanagh, Dinah Chancellor, Alma B. Ramirez, Jeanne F. Ramos, Kay Sammons, Jane F. Schielack, William Tate, Mary Thompson, and John A. Van de Walle. *Scott Foresman-Addison Wesley Mathematics: Grade 1.* Glenview, IL: Pearson Scott Foresman, 2005a.

Charles, Randall, Warren Crown, Francis (Skip) Fennell, Janet H. Caldwell, Mary Cavanagh, Dinah Chancellor, Alma B. Ramirez, Jeanne F. Ramos, Kay Sammons, Jane F. Schielack, William Tate, Mary Thompson, and John A. Van de Walle. *Scott Foresman-Addison Wesley Mathematics: Grade 2.* Glenview, IL: Pearson Scott Foresman, 2005b.

Clements, D.H., and J. Sarama. "Learning and Teaching Early and Elementary Mathematics." In *Instructional Strategies for Improving Student Learning: Focus on Early Mathematics and Reading* (vol. 3 of *Psychological Perspectives on Contemporary Educational Issues)*, edited by J.S. Carlson and J.R. Levine. Charlotte, NC: Information Age Publishing, 2012.

Clements, D.H., and J. Sarama *Learning and Teaching Early Math: The Learning Trajectories Approach*. New York, NY: Routledge, 2009.

Clements, D.H., and J. Sarama. *Engaging Young Children in Mathematics: Standards for Early Childhood Mathematics Education*. Mahwah, NJ: Lawrence Earlbaum Associates, Inc., 2004.

Cobb, P., and K. McClain. "Supporting Students' Learning of Significant Mathematical Ideas." In *Learning for Life in the 21st Century: Sociocultural Perspectives on the Future of Education*, edited by G. Wells and G. Claxton. Oxford, U.K.: Blackwell, 2002, pp. 154–166.

Cohen, D.K., and H.C. Hill. "Instructional Policy and Classroom Performance: The Mathematics Reform in California." *Teachers College Record*, vol. 102, no. 2, 2000, pp. 294–343.

Darch, C., D.W. Carnine, and R. Gersten. "Explicit Instruction in Mathematics Problem Solving." *Journal of Educational Research,* vol. 77, no. 6, July–August 1984, pp. 351–359.

Denton, K., and J. West. "Children's Reading and Mathematics Achievement in Kindergarten and First Grade." Publication No. NCES 2002-125. Washington, DC: U.S. Department of Education, National Center for Education Statistics, 2002.

Erchul, William P., and Brian K. Martens. *School Consultation: Conceptual and Empirical Bases of Practice.* Second ed. New York: Springer, 2002.

Fraivillig, J.L., L.A. Murphy, and K.C. Fuson. "Advancing Children's Mathematical Thinking in *Everyday Mathematics* Classrooms." *Journal for Research in Mathematics Education*, vol. 30, no. 2, March 1999, pp. 148–170.

Fuchs, L.S., D. Fuchs, and K. Karns. (2001). Enhancing Kindergartners' Mathematical Development: Effects of Peer-Assisted Learning Strategies. *Elementary School Journal,* vol. 101, 2001, pp. 495–510.

Fuchs, L.S., D. Fuchs, R. Finelli, S.J. Courey, C.L. Hamlett, E.M. Sones, and S.K. Hope. "Teaching Third-Graders About Real-Life Mathematical Problem Solving: A Randomized Controlled Study." *The Elementary School Journal*, vol. 106, 2006, pp. 293–311.

Fuchs, L.S., D. Fuchs, C.L. Hamlett, and K. Karns. "Acquisition and Transfer Effects of Classwide Peer-Assisted Learning Strategies in Mathematics for Students with Varying Learning Histories." *School Psychology Review*, vol. 24, no. 4, 1995, pp. 604–620.

Fuchs, L.S., D. Fuchs, C.L Hamlett, N.B. Phillips, K. Karns, and S. Dutka. "Enhancing Students' Helping Behavior During Peer-Mediated Instruction with Conceptual Mathematical Explanations." *The Elementary School Journal*, vol. 97, no. 3, January 1997, pp. 223–249.

Fuson, Karen C. *Math Expressions: Grade 1.* Orlando, FL: Houghton Mifflin Harcourt Publishing Company, 2009a.

Fuson, Karen C. *Math Expressions: Grade 2*. Orlando, FL: Houghton Mifflin Harcourt Publishing Company, 2009b.

Gersten, R., S. Beckmann, B. Clarke, A. Foegen, L. Marsh, J.R. Star, and B. Witzel. "Assisting Students Struggling with Mathematics: Response to Intervention (RtI) for Elementary and Middle Schools." NCEE 2009-4060. Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education, 2009. Available at [http://ies.ed.gov/ncee/wwc/PracticeGuide.aspx?sid=2].

Good, T.L., D.A. Grouws, and H. Ebmeier. *Active Mathematics Teaching*. New York, NY: Longman, 1983.

Good, T.L., and J.E. Brophy. *Looking in Classrooms*. Ninth ed. Columbus, OH: Allyn and Bacon, 2004.

Greenwood, C.R., J.C. Delquadri, and R.V. Hall. "Longitudinal Effects of Classwide Peer Tutoring." *Journal of Educational Psychology*, vol. 81, no. 3, September 1989, pp. 371–383.

Guarino, C., S.G. Dieterle, A.E. Bargagliotti, and W.M. Mason. "What Can We Learn About Effective Early Mathematics Teaching? A Framework for Estimating Causal Effects Using Longitudinal Survey Data." *Journal of Research on Educational Effectiveness*, vol. 6, no. 2, 2013, pp. 164–198.

Henry, V.J., and R.S. Brown. "First-Grade Basic Facts: An Investigation into Teaching and Learning of an Accelerated, High-Demand Memorization Standard." *Journal for Research in Mathematics Education*, vol. 39, no. 2, March 2008, pp. 153–183.

Hickey, D.T., A.L. Moore, and J.W. Pellegrino. "The Motivational and Academic Consequences of Elementary Mathematics Environments: Do Constructivist Innovations and Reforms Make a Difference?" *American Educational Research Journal,* vol. 38, no. 3, Fall 2001, pp. 611–652.

Hiebert, J.C., and D.A. Grouws. "The Effects of Classroom Mathematics Teaching on Students' Learning." In *Second Handbook of Research on Mathematics Teaching and Learning*, Vol. 1 (pp. 371–404), edited by F. K. Lester, Jr. New York, NY: Information Age Publishing, 2007.

Hill, Heather, Stephen G. Schilling, and Deborah Loewenberg Ball. "Developing Measures of Teachers' Mathematics Knowledge for Teaching." *The Elementary School Journal,* vol. 105, no. 1, September 2004, pp. 11–30.

Hopkins, K.B., A.V. McGillicuddy-De Lisi, and R. De Lisi. "Student Gender and Teaching Methods as Sources of Variability in Children's Computational Arithmetic Performance." *Journal of Genetic Psychology*, vol. 158, no. 3, September 1997, pp. 333–345.

Johnson, D.W., and R.T. Johnson. "An Educational Psychology Success Story: Social Interdependence Theory and Cooperative Learning." *Educational Researcher*, vol. 38, no. 5, 2009, pp. 365–379.

Kamii, C., and A. Dominick. "To Teach or Not to Teach Algorithms." *Journal of Mathematical Behavior*, vol. 16, no. 1, 1997, pp. 51–61.

Lampert, M., and P. Cobb. "Communication and Language." In *A Research Companion to Principles and Standards for School Mathematics*, edited by J. Kilpatrick, W. G. Martin, and D. Schifter. Reston, VA: National Council of Teachers of Mathematics, 2003, pp. 237–249.

Larson, Nancy, and Margaret Heisserer. *Saxon Math K–4: Teacher's Resource Handbook.* Orlando, FL: Harcourt Achieve, Inc., 2008.

Larson, Nancy, and Saxon Publishers. *Saxon Math 1 Lesson Sampler.* Austin, TX: Harcourt Achieve, Inc., 2006.

Le, V., B. Stecher, J.R Lockwood, L. Hamilton, A. Robyn, V. Williams, and S.P. Klein. “Improving Mathematics and Science Education: A Longitudinal Investigation Between Reform-Oriented Instruction and Student Achievement.” Santa Monica, CA: RAND Corporation, 2006.

Lemov, Doug. *Teach Like a Champion: 49 Techniques that Put Students on the Path to College.* San Francisco: Jossey-Bass, 2010.

Lord, F.M. *Applications of Item Response Theory to Practical Testing Problems*. Hillsdale, NJ: Lawrence Erlbaum Publishers, 1980.

Montague, M., and A.K. Jitendra. “Research-Based Mathematics Instruction for Students with Learning Disabilities.” In *Towards Equity in Mathematics Education*, edited by H. Forgasz and F. Rivera. Springer Berlin Heidelberg, 2012.

Murata, A., and K.C. Fuson. “Teaching as Assisting Individual Constructive Paths Within an Interdependent Class Learning Zone: Japanese First Graders Learning to Add Using 10.” *Journal for Research in Mathematics Education*, vol. 37, no. 5, November 2006, pp. 421–456.

Muthukrishna, N., and J.G. Borkowski. “How Learning Contexts Facilitate Strategy Transfer.” *Applied Cognitive Psychology*, vol. 9, no. 5, October 1995, pp. 425–446.

Nastasi, B.K., and D.H. Clements. “Research on Cooperative Learning: Implications for Practice.” *School Psychology Review*, vol. 20, no. 1, 1991, pp. 110–131.

National Assessment Governing Board. “Mathematics Framework for the 1996 National Assessment of Educational Progress.” Washington, DC: Government Printing Office, 1996.

National Association for the Education of Young Children (NAEYC). “10 Effective DAP Teaching Strategies.” Available at [https://www.naeyc.org/dap/10-effective-dap-teaching-strategies]. Accessed August 31, 2013.

National Center for Education Statistics (NCES). “The Nation’s Report Card: Mathematics 2011.” Washington, DC: Institute of Education Sciences, U.S. Department of Education, 2011.

National Council of Teachers of Mathematics. “Curriculum Focal Points for Prekindergarten Through Grade 8 Mathematics: A Quest for Coherence.” Reston, VA: National Council of Teachers of Mathematics, 2006.

National Mathematics Advisory Panel. “Foundations for Success: The Final Report of the National Mathematics Advisory Panel.” Washington, DC: U.S. Department of Education, Office of Planning, Evaluation and Policy Development, 2008.

National Research Council. *On Evaluating Curricular Effectiveness: Judging the Quality of K-12 Mathematics Evaluations*. Washington, DC: Mathematical Sciences Education Board, Center for Education, Division of Behavioral and Social Sciences and Education, National Academies Press, 2004.

National Research Council. *Mathematics in Early Childhood: Learning Paths toward Excellence and Equity*. Washington, DC: National Academy Press, 2009.

Pianta, R.C., J. Belsky, N. Vadergrift, R. Houts, and F.J. Morrison. "Classroom Effects on Children's Achievement Trajectories in Elementary School." *American Educational Research Journal*, vol. 45, no. 2, 2008, pp. 365–397.

Pianta, Robert C., Karen M. LaParo, and Bridget K. Hamre. *Classroom Assessment Scoring System:
K–3 Version*. Charlottesville, VA: Center for Advanced Study of Teaching and Learning, 2006.

Pollack, J., S. Atkins-Burnett, D. Rock, and M. Weiss. "Early Childhood Longitudinal Study—Kindergarten Class of 1998–99 (ECLS-K), Psychometric Report for the Third Grade." Washington, DC: U.S. Department of Education, National Center for Education Statistics, 2005.

Puma, M., R. Olsen, S. Bell, and C. Price. "What to Do When Data Are Missing in Group Randomized Controlled Trials." NCEE 2009-0049. Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education, 2009.

Remillard, Janine T. "What Else Matters?" Presented in the Symposium, Curriculum Matters: Why are Some Elementary School Curricula More Effective? Philadelphia, PA: NCTM Research Presession, April 2012.

Resnick, Robert M., Glenn Sanislo, and Stephanie Oda. *The Complete K–12 Report: Market Facts and Segment Analyses.* Rockaway Park, NY: Education Market Research, 2010.

Rittle-Johnson, B. "Promoting Transfer: Effects of Self-Explanation and Direct Instruction." *Child Development,* vol. 77, no. 1, January–February 2006, pp. 1–15.

Rock, Donald A., and Judith M. Pollack. "Early Childhood Longitudinal Study—Kindergarten Class of 1998–99 (ECLS-K), Psychometric Report for Kindergarten Through First Grade." Washington, DC: U.S. Department of Education, National Center for Education Statistics, 2002.

Rohrer, D., and H. Pashler. "Recent Research on Human Learning Challenges Conventional Instructional Strategies." *Educational Researcher*, vol. 39, no. 5, 2010, pp. 406–412.

Rosenshine, B. "Principles of Instruction: Research-Based Strategies That All Teachers Should Know." *American Educator*, vol. 36, no. 1, 2012, pp. 12–19, 39.

Sarama, J., and D.H. Clements. *Early Childhood Mathematics Education Research: Learning Trajectories for Young Children*. New York, NY: Routledge, 2009.

Shulman, L.S. “Those Who Understand: Knowledge Growth in Teaching.” *Educational Researcher*, vol. 15, no. 2, February 1986, pp. 4–14.

Siegler, R.S. “How Does Change Occur: A Microgenetic Study of Number Conservation.” *Cognitive Psychology*, vol. 28, no. 3, June 1995, pp. 255–273.

Slavin, R.E., and N.L. Karweit. “Effects of Whole Class, Ability Grouped, and Individualized Instruction on Mathematics Achievement.” *American Educational Research Journal*, vol. 22, no. 3, Fall 1985, pp. 351–367.

Stein, M.K., R.A. Engle, M.S. Smith, and E.K. Hughes. “Orchestrating Productive Mathematical Discussions: Five Practices for Helping Teachers Move Beyond Show and Tell.” *Mathematical Thinking and Learning*, vol. 10, no. 4, October 2008, pp. 313–340.

Stein, M.K., J. Remillard, and M.S. Smith. “How Curriculum Influences Student Learning.” In *Second Handbook of Research on Mathematics Teaching and Learning*, Vol. 1, edited by F. K. Lester, Jr. New York, NY: Information Age Publishing, 2007, pp. 319–369.

Sylva, K., E. Melhuish, P. Sammons, I. Siraj-Blatchford, and B. Taggart. *The Effective Provision of Pre-School Education [EPPE] Project: A Longitudinal Study Funded by the DfEE (1997–2003)*. London: EPPE Project, Institute of Education, University of London, 2005.

Waxman, Hersh C., Roland G. Tharp, and R. Soleste Hilberg (eds.). *Observational Research in U.S. Classrooms: New Approaches for Understanding Cultural and Linguistic Diversity*. Cambridge, UK: Cambridge University Press, 2004.

Webb, N.M. “Microcomputer Learning in Small Groups: Cognitive Requirements and Group Processes.” *Journal of Educational Psychology*, vol. 76, no. 6, December 1984, pp. 1076–1088.

Wenglinsky, H. “Closing the Racial Achievement Gap: The Role of Reforming Instructional Practices.” *Education Policy Analysis Archives*, vol. 12, no. 64, November 2004.

Wilkinson, L.A., A. Martino, and G. Camilli. “Groups that Work: Social Factors in Elementary Students’ Mathematics Problem Solving.” In *Research on Learning and Instruction of Mathematics in Kindergarten and Primary School*, edited by J.E.H. Van Luit (pp. 75–105). Doetinchem, Netherlands: Graviant, 1994.

Wittenberg, Lucy, K. Economopoulos, V. Bastable, K.H. Bloomfield, K. Cochran, D. Earnest, A. Hollister, N. Horowitz, E. Leidl, M. Murray, Y. Oh, B.W. Perry, S.J. Russell, D. Schifter, and K. Sillman. *Investigations in Number, Data, and Space.* Second ed. Glenview, IL: Pearson Scott Foresman, 2008.

Wright, B. “What Number Knowledge is Possessed by Children Beginning the Kindergarten Year of School?” *Mathematics Education Research Journal*, vol. 3, no. 1, 1991, pp. 1–16.

ENDNOTES

[1] More details about this study design and the effects of the curricula on student math achievement are reported in Agodini et al. (2008, 2009, 2010, and 2013).

[2] On average, the observation lasted 70 minutes per day and included the math lesson and morning meeting.

[3] Information about inter-rater reliability is provided in the Appendix.

[4] We examined the relationship between student achievement and specific practices because we believe it is more useful for informing future studies of instructional practices than examining relationships with the three scales. Specifically, because the scales were constructed from data collected for the first time with our custom-made observation protocol, we recommend that future research examines whether the scales underlying our protocol are robust across other samples of teachers prior to using them in this type of analysis.

[5] We do not focus on the relationship between the items in the classroom environment scale and student achievement because we did not find that this scale is related to achievement while working on the curriculum impact analysis.

[6] The ECLS-K bridge study was conducted to ensure that item overlap between the ECLS-K, K-1, and ECLS-K 3rd-grade items was adequate to place student achievement in a longitudinal scale (Pollack et al. 2005).

[7] The present study has a relatively high proportion of children of low socioeconomic status, and test results for the study's fall 2006 1st-grade sample showed mean math ability slightly below that of national ECLS-K fall 1st graders, by about one-eighth of a standard deviation. The selection of items included in the 2nd-grade test accounted for these factors.

[8] The bivariate correlations between the items included in each of the three first-step analyses are reported in the Appendix, Tables A.7 through A.12.

[9] The bivariate correlations between the items included in the second-step analysis are reported in the Appendix, Tables A.13 and A.14.

[10] On average, the observation lasted 70 minutes per day and included the math lesson and morning meeting.

[11] The Appendix also presents results from separate analyses for each instructional practice, indicating how each practice is related to student achievement without taking into consideration the influence of other practices and without adjusting for student, teacher, and school characteristics. A comparison of these results with those presented in Tables 3 through 5 illustrates how the adjustments affect the simple correlations reported in the Appendix.

For more information on the full study, please visit:

http://ies.ed.gov/ncee/projects/evaluation/math_curricula.asp

To read the technical appendix, please visit:

http://ies.ed.gov/ncee/pubs/20134020/pdf/20134020_app.pdf

This brief was prepared and based on analysis for NCEE by Douglas H. Clements, Roberto Agodini, and Barbara Harris under contract with Mathematica Policy Research (contract number ED-04-CO-0112/0003). Douglas H. Clements is affiliated with the University of Denver Morgridge College of Education, James C. Kennedy Institute for Educational Success, and Roberto Agodini and Barbara Harris are affiliated with Mathematica Policy Research. The NCEE project officer was Audrey Pendleton.