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

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

In this Appendix, we provide details about the data used for the current study, the curricula used in the classrooms from which data were collected, and the current study's methodological approach.

A. Data and Curricula

The data used for this study come from a randomized controlled trial (RCT) of math curricula (Agodini et al. 2008, 2009, 2010, and 2013).

1. Study Participants

The study team identified and recruited 12 districts and 110 schools that voluntarily participated in the study. Among the sites, 39 schools from 4 districts began study participation during the 2006–2007 school year, and the remaining 71 schools from 8 districts began participation during the 2007–2008 school year. Additional details about the study's recruitment effort are available in Agodini et al. (2010).[1]

In Tables A.1, A.2, and A.3, we present characteristics of the participating schools, teachers, and students, respectively, which may be summarized as follows:

- **Schools.** Compared to all U.S. elementary schools, those included in the analyses have a higher schoolwide Title I eligibility rate and a higher fraction of minority students and students eligible for free or reduced-price meals. In addition, the schools included in the analyses are located in larger districts than the average U.S. district; the districts included in the analyses have an average of 95 elementary schools, compared with 6 elementary schools in the average U.S. district.
- **Teachers.** The average study teacher is about 41 years old and has about 13 years of teaching experience. More than 90 percent of teachers are female. All teachers hold a bachelor's degree (about 80 percent of which are in an education field), and about a third also have a master's degree or higher. Most teachers took at least one math education course, and more than half took at least one advanced math course, such as trigonometry, calculus, or statistics.
- **Students.** About two-thirds of students are Hispanic or non-Hispanic black, and about 12 percent are limited-English proficient (LEP) or English-language learners

[1] As explained in Agodini et al. (2010), one of the study's school recruitment goals was to recruit schools that had not previously used any one of the four study curricula to help 'level the playing field' for the four curricula. Therefore, a key component of curriculum implementation involved training teachers to use their assigned curriculum. The study team did not mandate any minimum or maximum level of training, and instead supported any level of training each publisher deemed appropriate. The publishers provided all teachers (and support staff such as coaches) initial training prior to the first day of school and ongoing training and support throughout the school year.

(ELL). Upon entering the study, study students had an average math achievement that was slightly lower than a nationally representative sample of students who took the same test administered by the study about seven years earlier. As described below, the study team administered the math assessment developed for the Early Childhood Longitudinal Study-Kindergarten Class (ECLS-K) of 1998–1999.

Study participants are not a statistically representative sample of schools, teachers, and students with the above characteristics. Therefore, policymakers and educators should carefully examine the characteristics of the study participants when considering whether the results are useful for their own decision making.

Table A.1. Characteristics of U.S. Elementary Schools and Study Schools

	U.S. Elementary Schools	All Study Schools
Title I-Eligible (percentage)[a]	71.4	76.2
Schoolwide Title I-Eligible (percentage)	43.5	56.9
Student Enrollment (average)		
1st grade	71	91
2nd grade	69	86
Students Eligible for Free or Reduced-Price Meals (percentage)	46.9	49.9
Student Gender (percentage)		
Male	51.8	51.7
Female	48.2	48.3
Student Race/Ethnicity (percentage)		
White	58.2	38.5
Non-Hispanic black	16.4	32.1
Hispanic	19.3	26.2
Asian	4.0	1.9
American Indian or Alaska Native	2.2	1.4
Sample Size	**54,960**	**110**

Source: Author calculations using the 2005–2006 and 2006–2007 Common Core of Data (CCD). The "U.S. Elementary Schools" calculations include elementary schools with at least one 1st- or 2nd-grade student.

Table A.2. Characteristics of Study Teachers (percentages unless stated otherwise)

	1st-Grade Teachers	2nd-Grade Teachers
Average Age	40.5	40.6
Female	95.0	97.0
Race		
White	87.5	81.6
Other	12.5	18.4
Hispanic	20.3	31.1
Average Years of Teaching Experience	12.4	12.8
Highest Degree Earned		
Bachelor's degree	50.4	58.8
Master's degree	44.4	31.2
Number of Advanced Math Courses Taken		
None	44.4	47.2
1 or 2	42.4	37.4
3 or more	13.2	15.4
Number of Math Education Courses Taken		
None	3.8	5.1
1 or 2	55.2	39.8
3 or more	41.0	55.1
Teacher Assessment Score		
Overall	-0.54	-0.58
Content knowledge	-0.78	-0.86
Pedagogical knowledge	-0.33	-0.33
Sample Size	**365**	**277**

Source: Author calculations using fall teacher survey data and the study-administered assessment of teacher math content and pedagogical knowledge.

Table A.3. Characteristics of Study Students (percentages unless stated otherwise)

	1st-Grade Students	2nd-Grade Students
Fall Score (average)	31.3	55.8
Age at Fall Test (average)	6.6	7.7
Female	49.5	47.2
Race/Ethnicity		
Hispanic	32.0	35.3
Non-Hispanic black	23.7	25.4
Other non-Hispanic	44.3	39.3
LEP or ELL	14.4	10.5
Has IEP or Receives Special Services	7.4	7.9
Days Between Start of School and Fall Test (average)	21	22
Days Between Fall and Spring Tests (average)	238	237
Sample Size	**3,841**	**2,876**

Source: Author calculations using school records and data from the fall 1st- and 2nd-grade ECLS-K math test administered by the study.

2. Curricula in the Study

Investigations in Number, Data, and Space (Investigations), published by Pearson Scott Foresman, is a K–5 curriculum developed by the Technical Education Research Centers (TERC) under a grant from the National Science Foundation (NSF). It is based on a student-centered instructional approach that emphasizes metacognition, including thinking about one's own reasoning and that of peers; communicating mathematics orally and through writing and drawings; and solving problems in multiple ways. Students tend to work in depth on a smaller number of problems and are encouraged to choose from a variety of concrete materials and appropriate technology to help them solve problems as a regular part of their everyday work. Teachers spend much of their time facilitating conversations among students, helping students express their thoughts, and guiding them to a deeper understanding of mathematical concepts.

Each grade level is organized into units that last two to five weeks and focus on the exploration of major mathematical ideas. Units may focus on a single subject or revolve around a couple of related subjects—for example, addition and subtraction. Within each unit, the curriculum is built on two or more investigations that offer different contexts in which students explore mathematical problems using hands-on and written activities and class discussion. Some investigations last two or three days; others may last more than a week.

Classroom activities vary by day and depend on the length and type of investigation. For example, during a one-week investigation, the teacher introduces the material to the class on the first day, often through large-group, hands-on activities. During the next two to three days, students work in pairs or small groups to explore the concept, focus on a small number of in-depth problems, and play mathematical games. Every day, the teacher and students discuss as a group what they worked on, what they learned, and the strategies used to solve problems. At the end of the final day of the investigation, the teacher and students discuss the work, allowing students to compare solutions and strengthen their understanding. Daily routines, which can occur during the lesson or at some other time of day, are recommended in each unit and provide computation and data analysis practice.

Math Expressions, published by Houghton Mifflin Harcourt, is a K–5 curriculum based on the research results of the Children's Math Worlds (CMW) project conducted by Dr. Karen C. Fuson of Northwestern University and funded by the NSF. It uses a combination of teacher-directed and student-centered instructional approaches. Key aspects include specified algorithms; use of math language, math drawings, and visual representations; an emphasis on in-depth, sustained learning of core grade-level concepts (rather than a spiral curriculum); and skill fluency. The curriculum encourages teachers to provide students with efficient and effective procedures while promoting natural solution methods.

In Math Expressions classrooms, each day begins with student-led routines that involve the calendar, money, a number chart, counting, and time. The math lesson often occurs later in the day and begins with a quick fluency activity. Afterward, the teacher provides instruction to the whole class, introducing new information and encouraging students to discuss and demonstrate mathematical ideas. The teacher fosters the discussion while introducing efficient procedures; visual learning supports help students link their knowledge to formal mathematical concepts. Students then practice the new skill or concept in pairs, small groups, or individually, using worksheets. Homework is assigned daily.

Saxon Math (Saxon) is published by Houghton Mifflin Harcourt, and its primary program is a K–4 curriculum based on a teacher-directed instructional approach with scripted lesson plans, which are intended to help teachers deliver consistent and clear instruction to students. The program uses a multisensory approach with explicit instruction, hands-on activities, mathematical conversations, and practice. Each lesson integrates the mathematical strands, which are spiraled throughout the school year, so that concepts are developed, reviewed, and practiced over time rather than taught during discrete periods of time, such as in chapters or units. The teacher gradually introduces new material each day through explicit instruction and modeling. Each lesson includes daily distributed practice of previously learned concepts and procedures. The curriculum uses frequent, cumulative assessments to help teachers monitor student progress.

The curriculum is organized into five daily activities: (1) morning routines, (2) fact practice, (3) an explicit lesson, (4) guided class practice, and (5) homework. The morning routine is a whole-class activity that reinforces previously learned skills, lays the foundation for new skill development, allows students to work on problems in real-world settings, and often involves a student leader. The other four activities typically occur later in the day. Fact practice can occur during the same time as the math lesson, or at any other time; students work on fluency of number facts, either orally or in writing, with the support of self-correcting materials, manipulatives, fact cards, or worksheets. The lesson begins with a whole-class activity in which the teacher explicitly teaches the new concept using manipulatives and worksheets or overhead masters. After the lesson, the teacher guides practice while students use worksheets. At the end of each lesson, the teacher asks a few students to summarize for the entire class what they learned that day. Homework is assigned daily, and every fifth day teachers should administer a written or oral assessment to students.

Scott Foresman-Addison Wesley Mathematics (SFAW), published by Pearson Scott Foresman, is a pre-K–6 basal curriculum based on a teacher-directed approach that focuses on developing math skills and understanding. SFAW uses a consistent daily lesson structure that includes explicit instruction on mathematics skills and concepts, and hands-on exploration with manipulatives and pictorial and abstract representations. Essential outcomes and conceptual understandings are clearly articulated to teachers and students, and lessons include questioning strategies to develop students' higher-order thinking skills. Frequent and ongoing assessments and diagnoses are designed with strategic interventions to meet the individual needs of students, measure student understanding, and help guide instruction.

SFAW's consistent daily lesson structure includes six activities: (1) brief review of previously learned material; (2) hands-on exploration of a new concept; (3) brief activity to activate prior knowledge and connect it to the new lesson; (4) explicit instruction on the new concept in a whole-group setting; (5) individual, pair, or small-group practice using a worksheet or manipulatives; and (6) a closure activity to check student understanding of the new concept using worksheets, journal prompts, or questioning. The curriculum includes options for differentiating instruction within the consistent daily structure.

3. Data Collection

The study team collected data at the teacher, classroom, and student levels. Teacher data included scores on the assessment of math content and pedagogical knowledge, as well as survey data on teacher background characteristics and teaching practices. At the classroom level, the study team collected data from class rosters and conducted its own observations. At the student level, data sources were math tests administered by the study team and demographic information obtained from school records.

Below, we provide information about the classroom observations and student assessments. The data collection instruments are included in the study's design report (Agodini et al. 2008) with additional details about the data collection efforts in Agodini et al. (2010).

Classroom observations. The study team developed the protocol for conducting classroom observations specifically for the study. In designing the protocol, the study team first reviewed the study's curricula in depth, observed trainings by the publishers, and observed classrooms outside of the study that were implementing the study curricula. We identified critical features of each curriculum and asked publishers for feedback on the accuracy of these features. Once confirmed, we used the critical features as a basis for developing the protocol.

Next, keeping the features of each study curriculum in mind, the team reviewed the literature to identify methods previously used for assessing curriculum implementation and quality of instruction. Per the results of the review, the team developed a protocol that uses both interactive coding (counting and coding that clearly define behaviors as they occur) and ratings completed at the end of the observation (with use of a Likert scale to rate the degree to which different behaviors or characteristics are evident). Combining these approaches allows an observer to focus on the teacher-student interactions that occur and captures information about the frequency of those clearly defined interactions, while also gathering information about how evident or characteristic different behaviors are in the classroom.

The protocol was designed to include practices that may occur in implementation of all the curricula, as well as practices that are considered distinctive of one or more curricula. For many practices, the curricula differ in the extent to which the practice is used; therefore, most items on the protocol asked observers to tally the frequency of the practices. 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.

For the tally items, observers were instructed to stop tallying practices once the practice was observed 21 times. This cut-off was established to limit the burden placed on observers and because 21 instances of a practice was considered sufficient by the study team for differentiating the study's curricula. The study team carefully selected the cut-off of 21 as a likely discriminator across curricula, based on their review of the curricula and the critical features documents that were reviewed by the publishers, which suggested that if a practice was evident at least 21 times, it was very characteristic in a classroom and no further differentiation was needed for the purposes of supporting the study's impact analyses (the original intention for the data).

The observation protocol included approximately 100 cross-curriculum items organized into 10 sections that measure aspects of math instruction from a variety of perspectives. In general, 3 sections measure teacher behaviors, 3 sections measure student activities and materials, 2 sections measure instruction that pertains to teachers and students, and one section measures the percentage of time students spent in various groups during the lesson. The final section includes items that measure classroom management and some aspects of instruction (such as differentiation and peer collaboration).

Members of the study team were trained in the use of the protocol. Team members watched several classroom videos and learned how to code behaviors, interactions, and activities in each section of the protocol. After coding each video, a master coder led a group discussion of the results to bring observers to consensus on how to code each item. Observers had to pass a certification test on the entire protocol before conducting observations in the field. To become certified, an observer had to code within one category of the master observer on 85 percent of the items in the protocol. While the large number of items on the protocol could be considered a limitation of the study data, all items except one (as described below) met the study's criteria for reliability.

When observations took place in the field, study team members observed all math instruction throughout the day, including any morning meeting or calendar time, the math lesson, and any subsequent math instruction, such as drills or activity at math centers. Observers worked with teachers in advance to schedule observations and asked teachers to identify all times during the observation day when students were engaged in math instruction. In some classrooms, observers remained in the classroom for a single block of time; in others, they entered and exited the classroom several times throughout the day. To ensure that they were prepared to make accurate observations, observers reviewed the lesson to be taught before entering the classroom and had a copy of the lesson at hand for reference during the observation.

In 1st-grade classrooms, all observations took place in spring (March and April). In 2nd-grade classrooms, observations were evenly distributed across three points in the school year: (1) fall (October–November), (2) winter (January–February), and (3) spring (March–April). The study team attempted to observe all classrooms in schools with seven or fewer classrooms. In schools with more than seven classrooms, the team randomly sampled seven classrooms for observation.

Inter-rater reliability. To assess item reliability, two observers simultaneously coded about 10 percent of classroom observations. During the reliability observations, a master coder and classroom observer sat in the same classroom and independently observed all math instruction during the day of the observation. The two observers completed and submitted the classroom observation protocol separately and did not change any responses, regardless of any similarities or differences in coding. The paired observations were assessed for reliability using the same methods used to certify observers during the observation training. Specifically, for each item and each paired observation, we determined whether the master coder and classroom observer agreed or disagreed. Then, for each item, we calculated the percentage of paired observations that had agreement. Percentage agreement within one value was calculated for categorical and continuous items. For example, during a paired observation, if one of the observers coded 2 for a particular categorical item and the other observer coded 3 for the item, the two observers' codings on the item were considered to be in agreement.[2] Exact agreement was calculated for dichotomous items. Items with inter-rater reliability below 75 percent were considered unreliable. All items except one (which was an optional item to indicate the fraction of class time that students used any curricula-specific activity) met the study's criteria for inter-rater reliability; the one unreliable item was not used in the analyses summarized in this report (see Agodini et al. 2010).

Student testing. The study team administered to students the math assessment developed for the ECLS-K. The ECLS-K assessment is an adaptive test, in that it is tailored to a student's achievement level. In particular, the test begins with a short, first-stage routing test that broadly measures each student's math achievement level. Depending on the student's score on the routing test, the student is then assigned to one of three longer second-stage tests: (1) an easy test, (2) a middle-difficulty test, or (3) a difficult test. Some items on the second-stage tests are identical across the second-stage tests. The Educational Testing Service (ETS)—which was a developer of the test—uses Item Response Theory (IRT) techniques (Lord 1980) to analyze patterns of correct and incorrect answers to place scores from the different forms on the same scale for purposes of comparison. The assessment includes both open-ended and multiple-choice questions that measure conceptual understanding, procedural knowledge, and problem solving in five content areas: (1) number sense, properties, and operations; (2) measurement; (3) geometry and spatial sense; (4) data analysis, statistics, and probability; and (5) patterns, algebra, and functions.

Before initiating student testing, trained field staff collected class rosters from each participating classroom, reviewed the rosters with teachers to confirm that all students enrolled in

---

[2] Continuous (tallied) items were converted to the following seven categories for reliability assessments: 0 (0 tallies), 1 (1–2 tallies), 2 (3–5 tallies), 3 (6–10 tallies), 4 (11–15 tallies), 5 (16–20 tallies), and 6 (21 or more tallies). Percentage agreement was calculated within one value of these constructed categories for continuous items.

the class were listed on the roster, deleted from the roster the names of any students no longer enrolled, and identified students with language or other barriers (including physical and cognitive) that would make them ineligible for testing. Using updated rosters, field staff used a protocol to select, at random, eligible students in each classroom for testing. The number of students sampled in each class was a function of the number of classrooms in the target grade levels and class size; as a result, both fall and spring tests were administered to an average of 30 eligible students at the target grade levels in each study school.

Fall tests were administered within four weeks of the first day of classes; spring tests were administered one to six weeks before the end of the academic year. The test schedule aimed to administer the fall test as close as possible to the beginning of the school year and the spring test as close as possible to the end of the school year.

B. Methodological Approach

Below, we describe how we constructed the analysis file as well as the approach for conducting the analyses.

1. Construction of the Analysis File

The analysis sample includes 1st- and 2nd-grade students during the first year of curriculum implementation, along with observation data for their teachers. 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.

The analysis file contains student- , teacher- , and school-level measures. Student measures include scores on the ECLS-K math assessment and demographics from school records, including student age, gender, and race/ethnicity); whether a student is LEP or an ELL; and whether a student had an IEP or received special services. In addition, the analysis file includes the number of days between the beginning of school and the fall assessment and the number of days between the fall and spring assessments.

Teacher measures include their scores on a baseline assessment of math content and pedagogical knowledge administered by the study team, demographics from a fall survey administered by the team, and class rosters. The demographics from the survey include teacher experience, education, race/ethnicity, and previous use of the assigned curriculum at the K–3 level. Class size was extracted from the rosters. Last, to measure the heterogeneity of student achievement in each classroom, the classroom variance and skewness of the fall student math score was computed.

Two school-level measures were obtained from the CCD: (1) the percentage of students eligible for free or reduced-price meals and (2) whether the school is a Title I school. The block into which each school was placed during the random assignment process also was included in the analysis file, as well as the curriculum assigned to the school.

Complete data were available for the school-level measures and the fall and spring student math test scores. However, a small fraction of data was missing for some of the other student-level measures and for each teacher-level measure. As described in Agodini et al. (2010), the study teams used model-based imputations to replace the small amount of missing data.

A sampling weight was developed for each student. In particular, students in each file were weighted up to the number of students eligible for testing in the fall of 1st grade, separately for each classroom. For example, if 20 students in a classroom were eligible for testing in the fall but only 12 could be tested in the fall and spring of their relevant follow-up period, each of the 12 students received a weight of 1.67 (20/12). The weight was not adjusted for testing nonresponse; given the extent of missing test data observed in the current study, other research shows that simply analyzing pre- and post-tested students is equivalent to using weighting and imputation techniques that adjust for nonresponse (Puma et al. 2009).

2. Two-Step Analytic Approach

As described in the report, we used a two-step process to identify the instructional practices related to student achievement, separately for 1st and 2nd grade.[3] The first step examined the relationship between each practice and student achievement, taking into consideration the influence of other practices in the same scale. We estimated three hierarchical linear models (HLM), one for each set of items in three of the scales (named "student centered," "teacher directed," and "peer collaboration") that emerged from the classroom observation data (Agodini et al. 2010 describe the process for constructing the scales):

1. **Items in the "student-centered instruction" scale.** This includes 14 items that measure instructional practices expected in a student-centered instructional setting, including the extent to which teachers build on student thinking and elicit metacognitive understanding. At the 1st-grade level, the bivariate correlations between the items range from -0.04 to 0.51 (Table A.7). At the 2nd-grade level, the bivariate correlations between the items range from -0.07 to 0.56 (Table A.8).
2. **Items in the "teacher-directed instruction" scale.** This includes 15 items that measure instructional practices expected in a teacher-directed instructional setting, including the frequency of math practice and use of representations. At the 1st-grade level, the bivariate correlations between the items range from -0.13 to 0.75 (Table A.9). At the 2nd-grade level, the bivariate correlations between the items range from 0.01 to 0.78 (Table A.10).
3. **Items in the "peer collaboration" scale.** This includes 10 items that measure student interactions during math instruction, including teacher encouragement of student interactions, types of interactions, use of game playing, and percentage of time spent in small groups or pairs. At the 1st-grade level, the bivariate correlations between the items range from -0.09 to 0.79 (Table A.11). At the 2nd-grade level, the bivariate correlations between the items range from -0.06 to 0.77 (Table A.12).

Three-level HLMs were used to account for the nested structure of the data (students clustered in classrooms and classrooms clustered in schools):

[3] At both the 1st- and 2nd-grade levels, 82 percent of the variation in student math achievement is attributable to within-classroom differences, 6 percent is due to differences across classrooms within schools, and the remainder (12 percent) is due to differences across schools.

(1) $Y_{ijk} = \alpha_{0jk} + \alpha_{1jk} X_{ijk} + \varepsilon_{ijk}$,

(2) $\alpha_{0jk} = \beta_{0k} + \beta_{1k} W_{jk} + \mu_{jk}$,

(3) $\beta_{0k} = \delta_1 + \delta_2 Z_k + \nu_k$.

The first (student-level) equation assumes that spring math achievement can differ across students—that is, Y_{ijk} equals spring math achievement of student *i* in classroom *j* in school *k*, and α_{0jk} equals average spring math achievement of all students in classroom *j*, adjusted for average student characteristics, as represented by X_{ijk}. The last term in the equation (ε_{ijk}) represents the difference between spring achievement of student *i* and average spring achievement of all students in classroom *j*. We assume that ε_{ijk} is normally distributed with the same variance across classrooms.

The second (classroom-level) equation assumes that average adjusted spring classroom achievement from the first equation (α_{0jk}) can differ across classrooms because of teachers' instructional practices and other teacher/classroom characteristics, as represented by W_{jk}. The last term in the equation (μ_{jk}) represents the difference between average spring achievement of classroom *j* and average spring achievement of all classrooms in school *k*. We assume that μ_{jk} is normally distributed with the same variance across schools.

The third (school-level) equation assumes that average adjusted spring school achievement from the second equation (β_{0k}) can differ across schools because of school characteristics, as represented by Z_k. The last term in the equation (ν_k) represents the difference between average spring achievement of school *k* and average spring achievement of all schools.

In addition to including (in the second level) the instructional practices in a particular scale, the HLMs also included the following student, teacher/classroom, and school characteristics in the various levels of the HLMs:

The first (student) level of the HLM included the following:

- **Fall math score**—Student scale score on the math assessment administered at the beginning of the school year.
- **Days before the fall math assessment**—The number of days between the beginning of school and the student's fall math assessment.
- **Age**—Student age at the time of the fall math assessment.
- **Days between assessments**—The number of days between the student's fall and spring math assessments.
- **Gender**—Indicator of whether the student is female.

- **Race/ethnicity**—Indicators of whether the student is (1) Hispanic or (2) non-Hispanic black. Non-Hispanic white students and non-Hispanic students of other races serve as the reference category.
- **LEP/ELL**—Student is limited English proficient or an English language learner.
- **IEP/special services**—Student has an individualized education plan or receives special services.

The second (classroom) level of the HLM included the following:

- **Education**—Teacher has a master's degree. Teachers who do not have a master's degree, all of whom have a bachelor's degree, serve as the reference category.
- **Experience**—Years of teaching experience before the start of the first school year of the study.
- **Race/ethnicity**—Indicators of whether the teacher is (1) Hispanic or (2) non-Hispanic black. Non-Hispanic white teachers and non-Hispanic teachers of other races serve as the reference category.
- **Teacher baseline math assessment**—Teacher's overall scale score on the assessment of math content and pedagogical knowledge, measured at baseline.
- **Prior use of the assigned curriculum**—Teacher use of their school's randomly assigned curriculum at the K–3 level before joining the study.
- **Class size**—The number of students in the classroom in the fall.
- **Variance of the fall math score for the classroom**—Calculated variance of the student scale score on the fall assessment for the classroom.[4]
- **Skewness of the fall math score for the classroom**—Calculated skewness of the student scale score on the fall assessment for the classroom.[5]
- **The following individual items, where each item is on a four-point Likert scale that indicates how evident (not at all, minimally, strongly, and extremely) the classroom characteristics are:**
  - Students are cooperative and attentive to the lesson
  - Teacher spends a lot of time managing behavior (reverse coded)
  - Student behavior disrupts the classroom (reverse coded)
  - Students are perfectly behaved
  - Teacher uses nonverbal methods to manage misbehaviors

---

[4] Variance of the fall math score for the classroom was included in the HLM to account for the heterogeneity of students in each class.

[5] Skewness of the fall math score for the classroom was included in the HLM to account for the types of students (lower or higher achievers) that primarily comprise each class.

- Class runs without disruption from student behavior
- Students appear excited by the lesson
- Students are actively engaged
- Students attended to the lesson in a passive way (reverse coded)
- Students are off task (reverse coded)
- Teacher and students have a warm, positive relationship
- Teacher has techniques for gaining class attention in less than 10 seconds
- Students spend little time waiting or transitioning
- Transitions are smooth and students get to work quickly
- Teacher spends a lot of time giving directions (reverse coded)
- Teacher has materials prepared and ready for students
- Class time is spent on understanding or practicing math
- Teacher is fluid in presentation
- Students appear familiar with the materials and procedures used
- Students are given the opportunity to think and respond
- In monitoring student work, teacher followed through to ensure understanding

The third (school) level of the HLM included the following:

- **Free or reduced-price meals eligibility**—The percentage of students in the school eligible for free or reduced-price meals.
- **Title I**—An indicator of whether the school was Title I.

At the 1st-grade level, holding constant these student, teacher/classroom, and school characteristics, the “student-centered,” “teacher-directed,” and “peer collaboration” instructional practices included in the first-step analyses explain an additional 2.0, 5.3, and 2.5 percent of the variation in classroom-level achievement, respectively. At the 2nd-grade level, the “student-centered,” “teacher-directed,” and “peer collaboration” instructional practices explain an additional 4.2, 0.7, and 3.6 percent of the variation in classroom-level achievement, respectively.

The second step estimated one HLM that regressed spring student achievement on the practices in the first-step HLMs that were statistically significant at the 0.10 level. As above, a three-level HLM was used to account for the nested structure of the data (students clustered in classrooms and classrooms clustered in schools), and the practices that were statistically significant (at the 0.10 level) in the first-step HLMs were included in the second (teacher) level of the HLM. The HLM also included the student, teacher/classroom, and school characteristics listed above in the various levels of the HLM. Holding constant these student, teacher/classroom, and school characteristics, the instructional practices included in the second-step analysis explain an additional 3.7 and 6.4 percent of the variation in 1st- and 2nd-grade classroom-level achievement, respectively. The bivariate correlations for the practices examined in the second step are provided in Tables A.13 and A.14.

We examined results from the second step, and those that were statistically significant at the 0.10 level are deemed to be associated with student achievement and thus summarized in the report. We use the 0.10 level of statistical significance rather than the conventional level of 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.

The coefficient on a particular instructional practice indicates how a one-unit increase in that practice is associated with student achievement, adjusted for other practices; student, teacher, and school characteristics; and aspects of the classroom environment. The results are expressed as percentile point increases in achievement for the average student. We ran all results separately for 1st- and 2nd-grade students.

As described in the report, 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.*** This item was coded when the teacher responded to student work or answers by specifically naming a strategy and telling the student to use it, rather than suggesting different strategies the student could use or asking the student what they might try.
- ***Higher percentage of math instructional time spent in a large group or whole-class setting.*** This item indicates the percentage of time when students were working at their desks or together in a group as the teacher or student leader guided the class through problems, explained a new concept, or led a review of previously covered material.

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.*** "Above level" refers to students for whom the work is too easy. This item was coded when curriculum was differentiated by providing different materials, different worksheets, or changing the level of the lesson (that is, using more challenging numbers).
- ***Number of representations that teachers demonstrate.*** "Representations" refer to the ways in which math was demonstrated, such as pictures, diagrams, graphs, equations, number lines, and tables. Observers checked "yes" for each type of representation used by the teacher.
- ***Teachers asking the class if it agrees with a student's answer.*** This item was tallied for each time the teacher asked for some type of agreement or disagreement from other students. Students could have responded verbally or nonverbally (such as thumbs up or thumbs down).
- ***Students help one another understand math concepts or procedures.*** This student-focused item was coded after all math instruction was observed on a single day to characterize all class interactions during math instruction. The coding was based on the frequency with which students helped one another. Teachers may have

encouraged students to help one another, but students must actually have helped one another to be captured in the coding for this item.

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

- ***Teachers eliciting multiple strategies or solutions.*** This item tallied the number of problems for which the teacher elicited (and received) at least two different solutions or strategies. If the teacher asked for additional solutions but continued to get the same one, the problem was not tallied here.
- ***Teachers prompting a student to guide practice or lead the class in a routine.*** This item captures whether the teacher prompted a student in a leadership role. The teacher could prompt the student to ask math questions or lead the class through a routine.
- ***Frequency of students asking one another questions about math.*** This item captures the frequency with which students ask a peer about a solution, to explain thinking, or justify an answer. There must be actual interaction between students, not simply the teacher directing students to interact.

In Tables A.4 through A.6 at the end of this Appendix, we present the unadjusted relationship between each instructional practice and student achievement. Specifically, we estimated a separate HLM for each practice, regressing student achievement on the practice without controls for student, teacher, and school characteristics. A comparison of the results with those presented in the report (Tables 2 through 4) illustrates how the adjustments affect the simple correlations in Tables A.4 through A.6, respectively.

3. Other Considerations

As Hiebert and Grouws (2007) explained, the effects of a practice may depend on the system in which it functions, thereby pointing to the need to conduct separate analyses for each curriculum subgroup included in the RCT that serves as the basis for this study. We did not conduct these subgroup analyses because statistical power is a limiting factor. In particular, the small number of teachers in each subgroup could lead to insignificant findings about practices that a larger sample of teachers could demonstrate as important to student achievement.

In addition, our results on the influence of manipulatives and representations are based on a single measure for each of these two practices: one measure indicates the number of types of manipulatives used at least once and the other indicates the number of types of representations used at least once. The table below lists the manipulatives and representations included in each measure. We did not include a separate variable for each type of manipulative and representation in the analyses because of insufficient variation in teacher use of these items.

Manipulatives and Representations That Were Combined in Analytic Measures

Materials/Manipulatives Used	Types of Representations
Pattern blocks	Drawing pictures/dots
Linking cubes	Diagrams
Geoboards	Graphs
Blank cubes	Vertical equations/number sentence
Coins	Horizontal equations/number sentence
Base 10 blocks	Make a three-dimensional model
Fingers	Look for a pattern
Calculators	Tables
Dot cubes (dice)	Break-aparts
Counters	Labels
Ten frames	Number line
Stair steps	Sticks/circles
Fact/flash cards	Tallies
Math boards	Hundred (or 120) chart
Wrap-ups	Other (specify)
Part-part board	
Standard measuring tools	
Other (specify)	

Table A.4. Practices in the Student-Centered Scale: Percentile Point Increase in Student Achievement Associated with a One-Unit Increase in Each Practice, Based on Separate HLMs for Each Practice and No Controls for Other Practices, and on Student, Teacher, Classroom, and School Characteristics

	1st Grade		2nd Grade	
Item	Estimate	*p*-value	Estimate	*p*-value
Teacher poses open-ended questions that have more than one correct answer (TALLY)	0.3	0.12	0.2	0.18
Number of problems for which the teacher elicits multiple strategies or solutions (TALLY)	0.9**	0.05	0.5	0.37
Teacher tells student the strategy to use in response to student work/answer (TALLY)	1.2**	0.02	0.5	0.31
Teacher elicits other students' questions about a student's response (TALLY)	-0.6	0.74	1.3	0.33
Teacher labels math strategy, problem, or concept in response to student work/answer (TALLY)	1.2***	0.01	0.3	0.47
Teacher repeats student answer in a neutral way with no indication of correctness (TALLY)	0.3	0.24	0.4	0.34
Teacher probes for reasoning or justification in response to student work/answer (TALLY)	0.3	0.16	0.6***	0.01
Teacher provides hint to students in response to student work/answer (TALLY)	0.1	0.55	0.4**	0.01
Teacher clarifies what student says or does in response to student work/answer (TALLY)	0.6	0.15	0.3	0.58
Teacher extends what student says or does in response to student work/answer (TALLY)	1.9***	0.01	-0.6	0.32
Teacher uses praise or makes positive comments focused on content (TALLY)	1.2***	0.00	0.6	0.13
Teacher highlights student work or solution to class (TALLY)	0.3	0.46	1.3***	0.01
Number of different types of visual or three-dimensional representations created by students (TALLY)	0.7	0.13	1.0*	0.09
Teacher differentiates curriculum for children who are above level (SCALE, 1–4)	2.7	0.15	4.4**	0.02

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

*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 A.5. Practices in the Teacher-Directed Scale: Percentile Point Increase in Student Achievement Associated with a One-Unit Increase in Each Practice, Based on Separate HLMs for Each Practice and No Controls for Other Practices, and on Student, Teacher, Classroom, and School Characteristics

	1st Grade		2nd Grade	
Item	Estimate	*p*-value	Estimate	*p*-value
Teacher asks close-ended questions (TALLY)	-0.2	0.56	0.2	0.44
Number of problems on which the teacher guides practice on problems (TALLY)	0.1	0.39	0.4**	0.02
Number of representations demonstrated by the teacher (TALLY)	0.3	0.11	0.2	0.44
Teacher indicates if correct without elaborating in response to student work/answer (TALLY)	0.0	0.94	0.5**	0.03
Teacher calls on other students until the correct answer is given (TALLY)	-0.7**	0.03	-0.3	0.32
Teacher asks class if it agrees or disagrees with a student’s response (TALLY)	0.8**	0.02	0.4	0.16
Teacher prompts student to guide practice or lead the class in a routine (YES/NO)	1.8	0.38	1.7	0.47
Students practice number facts or procedures (SCALE, 1–6)	-0.3	0.43	0.5	0.21
Students provide choral or group responses to questions (SCALE, 0–2)	-1.0	0.50	-0.2	0.92
Students rote count (orally or in writing) (YES/NO)	2.7	0.21	-1.0	0.66
Number of types of rote counting that occurred, by ones, twos, and so forth (TOTAL OF 8 TIMES)	1.0	0.13	1.7*	0.05
Number of practice problems focusing on review of previously learned material (TALLY)	0.3*	0.05	0.1	0.48
Number of materials used by children (TOTAL OF 11 ITEMS)	-0.6	0.55	0.1	0.91
Number of types of representations used during math, by the teacher or by students (TOTAL OF 7 ITEMS)	0.2	0.79	0.6	0.50
Percentage of math instructional time spent in large group (SCALE, 0–4)	0.4	0.70	-0.3	0.82

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

*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 A.6. Practices in the Peer Collaboration Scale: Percentile Point Increase in Student Achievement Associated with a One-Unit Increase in Each Practice, Based on Separate HLMs for Each Practice and No Controls for Other Practices, and on Student, Teacher, Classroom, and School Characteristics

	1st Grade		2nd Grade	
Item	Estimate	*p*-value	Estimate	*p*-value
Teacher demonstrates how to play a game (YES/NO)	-2.0	0.43	2.1	0.52
Teacher directs or encourages students to help one another with math (YES/NO)	1.8	0.33	1.2	0.59
Students play math games (SCALE, 0–6)	-0.4	0.28	0.5	0.38
Students ask peers questions about math (SCALE, 0–2)	2.9*	0.09	2.8	0.15
Students discuss math strategies or solutions with partner or small group (SCALE, 0–2)	2.5*	0.09	1.7	0.36
Percentage of math instructional time spent in small group (SCALE, 0–4)	0.4	0.74	-0.1	0.93
Percentage of math instructional time spent in pairs (SCALE, 0–4)	-0.1	0.93	-0.6	0.62
Teacher encourages students to help one another understand math (SCALE, 1–4)	0.7	0.52	3.1**	0.02
Students help one another understand math concepts or procedures (SCALE, 1–4)	1.0	0.43	3.0**	0.02
Peer-to-peer interaction about math occurs (SCALE, 1–4)	1.2	0.34	1.7	0.28

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

*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 A.7. Bivariate Correlations, Practices Within the Student-Centered Scale: 1st-Grade (N = 363)

Item	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Teacher poses open-ended questions that have more than one correct answer (TALLY)	1.00	0.46	0.00	0.09	0.03	0.42	0.31	0.07	0.24	0.20	0.09	0.14	0.15	0.14
2. Number of problems for which the teacher elicits multiple strategies or solutions (TALLY)		1.00	0.06	0.20	0.16	0.19	0.20	0.06	0.16	0.14	0.21	0.25	0.19	0.16
3. Teacher tells student the strategy to use in response to student work/answer (TALLY)			1.00	0.13	0.21	0.16	0.03	0.14	0.10	0.03	0.26	-0.03	0.17	0.07
4. Teacher elicits other students' questions about a student's response (TALLY)				1.00	0.09	0.12	0.12	0.07	0.09	0.13	0.10	0.07	0.04	0.05
5. Teacher labels math strategy, problem, or concept in response to student work/answer (TALLY)					1.00	0.07	0.18	0.25	0.25	0.22	0.37	0.20	0.31	0.04
6. Teacher repeats student answer in a neutral way with no indication of correctness (TALLY)						1.00	0.23	0.19	0.13	0.18	0.19	0.10	0.16	0.19
7. Teacher probes for reasoning or justification in response to student work/answer (TALLY)							1.00	0.31	0.52	0.30	0.09	0.25	0.22	0.06
8. Teacher provides hint to students in response to student work/answer (TALLY)								1.00	0.23	0.08	0.10	0.09	0.17	-0.04
9. Teacher clarifies what student says or does in response to student work/answer (TALLY)									1.00	0.40	0.10	0.27	0.22	0.07
10. Teacher extends what student says or does in response to student work/answer (TALLY)										1.00	0.17	0.19	0.08	0.06
11. Teacher uses praise or makes positive comments focused on content (TALLY)											1.00	0.05	0.20	0.21
12. Teacher highlights student work or solution to class (TALLY)												1.00	0.19	0.00
13. Number of different types of visual or three-dimensional representations created by students (TALLY)													1.00	0.08
14. Teacher differentiates curriculum for children who are above grade level (SCALE, 1–4)														1.00

Note: Pearson correlation coefficients calculated by the authors.

Table A.8. Bivariate Correlations, Practices Within the Student-Centered Scale: 2nd-Grade (N = 269)

Item	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Teacher poses open-ended questions that have more than one correct answer (TALLY)	1.00	0.56	0.16	0.11	0.03	0.32	0.27	0.11	0.27	0.08	0.00	0.12	0.19	0.11
2. Number of problems for which the teacher elicits multiple strategies or solutions (TALLY)		1.00	-0.01	0.24	0.07	0.07	0.22	0.15	0.24	0.03	0.07	0.23	0.22	0.04
3. Teacher tells student the strategy to use in response to student work/answer (TALLY)			1.00	0.08	0.21	0.17	0.30	0.33	0.08	0.16	0.02	0.15	0.08	0.15
4. Teacher elicits other students' questions about a student's response (TALLY)				1.00	0.02	0.09	0.20	0.14	0.21	0.15	0.14	0.05	0.23	0.10
5. Teacher labels math strategy, problem, or concept in response to student work/answer (TALLY)					1.00	0.05	0.13	0.25	0.05	0.21	0.16	0.10	0.18	0.04
6. Teacher repeats student answer in a neutral way with no indication of correctness (TALLY)						1.00	0.23	0.18	0.10	0.08	-0.07	0.14	0.12	0.11
7. Teacher probes for reasoning or justification in response to student work/answer (TALLY)							1.00	0.25	0.50	0.10	0.15	0.29	0.13	0.14
8. Teacher provides hint to students in response to student work/answer (TALLY)								1.00	0.14	0.05	0.02	0.06	0.29	0.17
9. Teacher clarifies what student says or does in response to student work/answer (TALLY)									1.00	0.14	0.19	0.18	0.12	0.19
10. Teacher extends what student says or does in response to student work/answer (TALLY)										1.00	0.23	0.15	0.02	0.07
11. Teacher uses praise or makes positive comments focused on content (TALLY)											1.00	0.11	0.01	-0.02
12. Teacher highlights student work or solution to class (TALLY)												1.00	0.09	0.06
13. Number of different types of visual or three-dimensional representations created by students (TALLY)													1.00	0.10
14. Teacher differentiates curriculum for children who are above grade level (SCALE, 1–4)														1.00

Note: Pearson correlation coefficients calculated by the authors.

Table A.9. Bivariate Correlations, Practices Within the Teacher-Directed Scale: 1st-Grade (N = 362)

Item	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. Teacher asks close-ended questions (TALLY)	1.00	0.16	0.17	0.61	0.11	0.10	0.05	0.18	0.25	0.07	0.15	0.05	-0.08	0.06	0.28
2. Number of problems on which the teacher guides practice on problems (TALLY)		1.00	0.33	0.14	0.10	0.11	0.03	0.18	0.10	0.21	0.24	0.30	0.10	0.12	0.24
3. Number of representations demonstrated by the teacher (TALLY)			1.00	0.19	0.16	0.27	0.24	0.23	0.09	0.37	0.48	0.36	0.29	0.37	0.20
4. Teacher indicates if correct without elaborating in response to student work/answer (TALLY)				1.00	0.11	0.03	0.06	0.10	0.08	0.16	0.19	0.19	-0.06	0.05	0.19
5. Teacher calls on other students until the correct answer is given (TALLY)					1.00	0.10	-0.03	0.01	0.03	0.15	0.17	0.07	-0.03	0.08	0.13
6. Teacher asks class if it agrees or disagrees with a student's response (TALLY)						1.00	0.14	0.14	0.18	0.13	0.24	0.20	0.05	0.12	0.11
7. Teacher prompts student to guide practice or lead the class in a routine (YES/NO)							1.00	0.03	0.03	0.25	0.23	0.04	-0.01	0.08	0.07
8. Students practice number facts or procedures (SCALE, 1–6)								1.00	0.22	0.08	0.16	0.29	0.04	0.14	0.17
9. Students provide choral or group responses to questions (SCALE, 0–2)									1.00	-0.02	0.10	0.05	-0.01	0.08	0.28
10. Students rote count (orally or in writing) (YES/NO)										1.00	0.75	0.18	0.09	0.13	0.12
11. Number of types of rote counting that occurred, by ones, twos, and so forth (TOTAL OF 8 TIMES)											1.00	0.29	0.17	0.20	0.11
12. Number of practice problems focusing on review of previously learned material (TALLY)												1.00	0.11	0.35	0.09
13. Number of materials used by children (TOTAL OF 11 ITEMS)													1.00	0.24	-0.13
14. Number of types of representations used during math, by the teacher or by students (TOTAL OF 7 ITEMS)														1.00	0.13
15. Percentage of math instructional time spent in large group (SCALE, 0–4)															1.00

Note: Pearson correlation coefficients calculated by the authors.

Table A.10. Bivariate Correlations, Practices Within the Teacher-Directed Scale: 2nd-Grade (N = 269)

Item	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. Teacher asks close-ended questions (TALLY)	1.00	0.20	0.26	0.61	0.17	0.12	0.19	0.07	0.23	0.16	0.21	0.20	0.01	0.15	0.34
2. Number of problems on which the teacher guides practice on problems (TALLY)		1.00	0.19	0.29	0.26	0.09	0.11	0.20	0.24	0.14	0.20	0.25	0.02	0.13	0.33
3. Number of representations demonstrated by the teacher (TALLY)			1.00	0.22	0.24	0.30	0.51	0.25	0.16	0.49	0.62	0.50	0.32	0.51	0.32
4. Teacher indicates if correct without elaborating in response to student work/answer (TALLY)				1.00	0.23	0.16	0.17	0.08	0.06	0.15	0.14	0.26	0.05	0.13	0.22
5. Teacher calls on other students until the correct answer is given (TALLY)					1.00	0.18	0.20	0.12	0.12	0.20	0.20	0.22	0.09	0.19	0.27
6. Teacher asks class if it agrees or disagrees with a student's response (TALLY)						1.00	0.26	0.17	0.27	0.20	0.26	0.25	0.17	0.23	0.18
7. Teacher prompts student to guide practice or lead the class in a routine (YES/NO)							1.00	0.13	0.13	0.42	0.44	0.25	0.12	0.19	0.22
8. Students practice number facts or procedures (SCALE, 1–6)								1.00	0.19	0.25	0.36	0.39	0.13	0.12	0.20
9. Students provide choral or group responses to questions (SCALE, 0–2)									1.00	0.17	0.22	0.06	0.04	0.07	0.25
10. Students rote count (orally or in writing) (YES/NO)										1.00	0.78	0.33	0.19	0.22	0.28
11. Number of types of rote counting that occurred, by ones, twos, and so forth (TOTAL OF 8 TIMES)											1.00	0.45	0.22	0.31	0.30
12. Number of practice problems focusing on review of previously learned material (TALLY)												1.00	0.25	0.33	0.27
13. Number of materials used by children (TOTAL OF 11 ITEMS)													1.00	0.24	0.20
14. Number of types of representations used during math, by the teacher or by students (TOTAL OF 7 ITEMS)														1.00	0.10
15. Percentage of math instructional time spent in large group (SCALE, 0–4)															1.00

Note: Pearson correlation coefficients calculated by the authors.

Table A.11. Bivariate Correlations, Practices Within the Peer Collaboration Scale: 1st-Grade (N = 357)

Item	1	2	3	4	5	6	7	8	9	10
1. Teacher demonstrates how to play a game (YES/NO)	1.00	0.15	0.79	0.13	0.27	0.19	0.36	0.25	0.29	0.42
2. Teacher directs or encourages students to help one another with math (YES/NO)		1.00	0.18	0.38	0.40	0.19	0.23	0.64	0.48	0.42
3. Students play math games (SCALE, 0–6)			1.00	0.14	0.33	0.16	0.45	0.26	0.33	0.43
4. Students ask peers questions about math (SCALE, 0–2)				1.00	0.59	0.24	0.23	0.52	0.57	0.54
5. Students discuss math strategies or solutions with partner or small group (SCALE, 0–2)					1.00	0.18	0.45	0.62	0.71	0.61
6. Percentage of math instructional time spent in small group (SCALE, 0–4)						1.00	-0.09	0.16	0.17	0.26
7. Percentage of math instructional time spent in pairs (SCALE, 0–4)							1.00	0.34	0.47	0.58
8. Teacher encourages students to help one another understand math (SCALE, 1–4)								1.00	0.73	0.60
9. Students help one another understand math concepts or procedures (SCALE, 1–4)									1.00	0.73
10. Peer-to-peer interaction about math occurs (SCALE, 1–4)										1.00

Note: Pearson correlation coefficients calculated by the authors.

Table A.12. Bivariate Correlations, Practices Within the Peer Collaboration Scale: 2nd-Grade (N = 265)

Item	1	2	3	4	5	6	7	8	9	10
1. Teacher demonstrates how to play a game (YES/NO)	1.00	0.07	0.77	0.11	0.13	0.22	0.24	0.07	0.11	0.25
2. Teacher directs or encourages students to help one another with math (YES/NO)		1.00	0.08	0.47	0.39	0.12	0.36	0.62	0.46	0.40
3. Students play math games (SCALE, 0–6)			1.00	0.08	0.16	0.21	0.31	0.09	0.12	0.29
4. Students ask peers questions about math (SCALE, 0–2)				1.00	0.52	0.26	0.21	0.63	0.62	0.56
5. Students discuss math strategies or solutions with partner or small group (SCALE, 0–2)					1.00	0.21	0.48	0.49	0.54	0.65
6. Percentage of math instructional time spent in small group (SCALE, 0–4)						1.00	-0.06	0.17	0.23	0.28
7. Percentage of math instructional time spent in pairs (SCALE, 0–4)							1.00	0.35	0.35	0.56
8. Teacher encourages students to help one another understand math (SCALE, 1–4)								1.00	0.75	0.61
9. Students help one another understand math concepts or procedures (SCALE, 1–4)									1.00	0.69
10. Peer-to-peer interaction about math occurs (SCALE, 1–4)										1.00

Note: Pearson correlation coefficients calculated by the authors.

Table A.13. Bivariate Correlations, 1st-Grade Practices Examined in Step 2 of the Analysis (N = 362)

Item	1	2	3	4	5
1. Students help one another understand math concepts or procedures (SCALE, 1–4)	1.00	-0.06	0.08	-0.12	-0.24
2. Teacher tells student the strategy to use in response to student work/answer (TALLY)		1.00	0.14	0.01	-0.04
3. Teacher provides hint to students in response to student work/answer (TALLY)			1.00	0.09	0.03
4. Teacher calls on other students until the correct answer is given (TALLY)				1.00	0.13
5. Percentage of math instructional time spent in large group (SCALE, 0–4)					1.00

Note: Pearson correlation coefficients calculated by the authors.

Table A.14. Bivariate Correlations, 2nd-Grade Practices Examined in Step 2 of the Analysis (N = 269)

Item	1	2	3	4	5	6	7	8	9
1. Students ask peers questions about math (SCALE, 0–2)	1.00	0.62	0.07	0.02	0.04	0.26	-0.11	0.06	0.10
2. Students help one another understand math concepts or procedures (SCALE, 1–4)		1.00	0.08	0.07	0.11	0.18	-0.02	0.06	0.06
3. Teacher poses open-ended questions that have more than one correct answer (TALLY)			1.00	0.56	0.32	0.11	0.13	0.05	-0.04
4. Number of problems for which the teacher elicits multiple strategies or solutions (TALLY)				1.00	0.07	0.04	0.17	0.04	0.06
5. Teacher repeats student answer in a neutral way with no indication of correctness (TALLY)					1.00	0.11	-0.09	0.04	-0.03
6. Teacher differentiates curriculum for children who are above grade level (SCALE, 1–4)						1.00	-0.03	-0.01	0.11
7. Number of representations demonstrated by the teacher (TALLY)							1.00	0.30	0.51
8. Teacher asks class if it agrees or disagrees with a student's response (TALLY)								1.00	0.26
9. Teacher prompts student to guide practice or lead the class in a routine (YES/NO)									1.00

Note: Pearson correlation coefficients calculated by the authors.

For more information on the full study, please visit:

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

To read the evaluation brief, please visit:

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

ies NATIONAL CENTER FOR EDUCATION EVALUATION AND REGIONAL ASSISTANCE

Institute of Education Sciences

MATHEMATICA Policy Research

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.