J Educ Psychol. Author manuscript; available in PMC 2010 August 1.

Published in final edited form as:
J Educ Psychol. 2010 May ; 102(2): 407–417. doi:10.1037/a0017762.

Instruction, Teacher–Student Relations, and Math Achievement Trajectories in Elementary School

Robert Crosnoe,
Department of Sociology and Population Research Center, University of Texas at Austin

Fred Morrison,
Department of Psychology, University of Michigan

Margaret Burchinal,
Frank Porter Graham Child Development, University of North Carolina at Chapel Hill

Robert Pianta,
Curry School of Education, University of Virginia

Daniel Keating,
Department of Psychology, University of Michigan

Sarah L. Friedman,
CNA Corporation, Alexandria, VA

K. Alison Clarke-Stewart, and
Department of Psychology and Social Behavior, University of California, Irvine

The Eunice Kennedy Shriver National Institute of Child Health and Human Development Early Child Care Research Network

Abstract

Children enter elementary school with widely different skill levels in core subjects. Whether because of differences in aptitude or in preparedness, these initial skill differences often translate into systematic disparities in achievement over time. How can teachers reduce these disparities? Three possibilities are to offer basic skills training, to expose students to higher order instruction, or to provide socioemotional support. Repeated measures analyses of longitudinal data from the Eunice Kennedy Shriver National Institute of Child Health and Human Development Study of Early Child Care and Youth Development revealed that children with low, average, or high math skills prior to elementary school followed different but parallel trajectories of math achievement up through fifth grade. When enrolled in classes with inference-based instruction, however, the initially least skilled children narrowed the achievement gap as long as they did not have conflictual relations with their teachers. They did not make this kind of progress if they were in classes focused exclusively on basic skills instruction or if they were in inference-focused classes but had conflictual relations with teachers.

Keywords

elementary school; instruction; teacher–student relations; math

Correspondence concerning this article should be addressed to Robert Crosnoe, Department of Sociology and Population Research Center, University of Texas at Austin, 1 University Station A1700, Austin, TX 78712-1088. crosnoe@austin.utexas.edu.

The beginning of elementary school is often referred to as the *starting gate*, meaning that it is the threshold of a long career in the educational system that has profound implications for the life course (Lee & Burkham, 2002). Yet, this term is something of a misnomer. Although this period signals the beginning of formal schooling, it is not the beginning of the academic process. Instead, children can be engaged in learning—sometimes in quite formal, organized ways—for years leading up to school entry: in their homes, in preschools, in other settings. Because these early learning activities are not institutionalized, however, children hit the official starting gate of their educational careers with widely different sets of skills and knowledge (Eunice Kennedy Shriver National Institute of Child Health and Human Development [NICHD] Early Child Care Research Network, 2005b; Raver, Gershoff, & Aber, 2007). Theoretically at least, formal schooling is supposed to reduce these initial differences, providing children who lack early learning opportunities with the kinds of instruction and stimulation that they need to catch up to their peers. Unfortunately, the opposite is much more likely to happen, because formal schooling magnifies the advantages that some children bring with them into the system so that initial disparities compound over time (Alexander & Entwisle, 1988).

Curriculum seems key to both evening out and increasing the skills differences children bring into school (Hamre & Pianta, 2005). The norm now is for children with less developed skills to be offered a watered-down curriculum deemed more aptitude appropriate, whereas their more advanced peers are exposed to more challenging and ultimately more rewarding classes. This situation, then, stratifies learning opportunities in elementary school based on differences in learning opportunities prior to elementary school. In contrast, a common curriculum equally exposes all children to high-end instruction. This would not affect the already advanced children, but it would likely foster more learning among the less advanced children, thereby promoting convergence in achievement trajectories over time (Connor, Morrison, & Katch, 2004; Xue & Meisels, 2004). Yet, this approach also poses a major risk. Challenging course work absent adequate instrumental and social support for meeting those challenges could be a recipe for failure for children who begin school with a more limited skill set. Thus, the answer is not simply more challenging classes but, instead, classes that are advanced and supportive. Such support may entail a supplementary focus on foundational skills that are not available to the children entering with fewer skills and/or positive emotional responses from teachers (Connor, Morrison, Fishman, Schatschneider, & Underwood, 2007; Griffin, 2004). The intersection of emotional and instructional quality matters (Greenberg et al., 2003; Howes, 2000).

This study draws on the NICHD Study of Early Child Care and Youth Development (SECCYD) to explore the possibility that breaking the link between entry-level academic skills and classroom pedagogy will allow less school-ready children to catch up with their more advanced peers over time. Specifically, we examined whether such catch up was more likely to occur among children in cognitively challenging classrooms—in which instruction emphasized conceptual understanding and analyses, not just learning and reciting factual material—led by teachers with whom children had good relations than among children without one or both of these classroom/teacher features. Such research informs both theory and policy by putting forward the balance between the formal, instructional and informal, socioemotional aspects of schooling as a tool for promoting equity in the American educational system.

Inequality in Early Education

Long before children enter elementary school, their parents can create a variety of learning opportunities for them (Pianta & Cox, 1999). For example, parents can construct cognitively stimulating home environments with shared reading time and developmentally appropriate learning materials. They can also enroll children in preschools with organized curricula led by trained personnel and/or sign them up for lessons and activities (Magnuson, Meyers, Ruhm, & Waldfogel, 2004; NICHD Early Child Care Research Network, 2005b; Raver et al., 2007; Waldfogel, 2006). Children's exposure to such stimulation, however, is largely predicated on their parents' economic resources, social contexts, and personal motives. Consequently, variability in learning experiences prior to school entry translates into equal variability in school readiness above and beyond existing differences in intellectual competencies (Crosnoe, 2006; Sadowski, 2006).[1]

As a result of these different early learning experiences, some children enter school with low skill levels in core curricula, like math, and others enter with much more advanced skills. These skill differences may reflect variations in aptitude, preparedness, or both. Regardless, such differences in school readiness then drive class placement, teacher expectations, and other factors that, over time, widen initial disparities (Entwisle & Alexander, 2002). As an example, consider two children who start school with relatively similar aptitudes but very different levels of exposure to math. On the basis of these differences, the more prepared child would likely be assigned to a more advanced math class, and his or her peer would likely be assigned to a more basic skills class. Because the former class probably has a better teacher with higher expectations for students and a more rigorous, higher quality lesson plan that builds on existing skill sets, the first child is likely to learn more over the course of the year than the second. These differences in achievement, in part predicated on initial differences in preparedness, would widen the skill gap and send these two children in increasingly different directions as years pass (Entwisle & Alexander, 2002; Farkas, 1996; Pianta et al., 2005; Rimm-Kaufman, La Paro, Pianta, & Downer, 2005; Sadowski, 2006).

A major theme of educational research across disciplines has been the search for mechanisms that, when targeted by policy intervention, might break this cycle of cumulative disadvantage and advantage (Ceci & Papierno, 2005). Often, aspects of formal classroom organization are the focus of this search. The goal is to identify classroom features that help children with learning lags catch up with their peers. In this study, we pursue this goal by looking into various instructional processes in elementary school classrooms, paying special attention to their interplay with teacher–student relations.

Elementary School Classrooms and Learning Trajectories

A growing multidisciplinary literature has highlighted the potential value of a common curriculum. When students, regardless of aptitude level, take the same classes, their opportunities to learn are more equitably distributed. The more advanced students get the instruction that they would normally receive, but the less advanced students get better instruction than they would normally receive. The higher average achievement of the latter in such situations suggests that they rise to meet challenges (Bryk, Lee, & Holland, 1993).

This idea, largely built on studies of private high schools in the United States, is also valuable for thinking about ways to redesign the elementary school experience. When implemented correctly, higher order instruction (e.g., teaching strategies that highlight deduction and critical thinking) in conjunction with skill acquisition tends to promote achievement more than a rote basic skills approach, especially among students who initially are academically at risk (Cameron, Connor, & Morrison, 2005; Connor et al., 2007; Griffin, 2004; Hamre & Pianta, 2005; Xue & Meisels, 2004). Consequently, the tight coupling of higher order instruction with more skilled students and of exclusively basic skills instruction with less skilled students likely makes achievement disparities worse, not better.

[1]This stratification system is closely related to racial, ethnic, and socioeconomic inequality.

Alternatively, exposing all children to similar levels of higher order instruction may close, rather than widen, achievement disparities. More to the point, children who have low skill levels in a subject like math at the start of school would likely make up ground over time if they enjoyed the same instructional resources as their peers who, at entry level, were more skilled. If, on the other hand, they are relegated to less demanding courses because of their perceived lack of aptitude, they will likely fall farther behind.

The linchpin here, of course, is correct implementation. Haphazardly exposing relatively less prepared or able children to the same kind of challenging, complex activities and instruction as their better prepared peers could be counterproductive. They might be overwhelmed or simply not have the necessary foundation, in which case common exposure would add to the cumulative disadvantage process it is intended to correct. Thus, care must be taken to make this experience less jarring. For example, initially low-skilled children typically gain more reading ability at the start of elementary school when their teachers pursue higher order instructional strategies along with basic skill supplements. Their initially high-skilled peers do not need those supplements. Eventually, the two groups move toward convergence (Connor et al., 2004).

Another way to approach the issue of correct implementation is to think about the context of higher order instruction. Such instruction is challenging, with setbacks and stumbles that create tension. To successfully navigate a potentially at-risk child through such an experience, a teacher would need to attend to her or his psychological well-being. A teacher who can provide encouragement, emotional support, and comfort to a student is likely to be better able to make that higher order approach work. In contrast, a teacher–student relationship fraught with conflict (e.g., a teacher expresses frustration with a struggling student, a student expresses resentment to a teacher challenging him or her) is not conducive to success (Pianta, Steinberg, & Rollins, 1995). Indeed, having conflictual relations with teachers in a cognitively challenging classroom might be even worse for achievement than the standard basic skills classrooms to which most slow starters are assigned.

Worth noting is that exposing all children, regardless of initial skill level, to challenging classrooms led by supportive teachers cannot completely close subsequent achievement gaps for the basic reason that those initial differences in skill are not simply a function of differences in preparedness and experience. They also likely reflect naturally occurring differences in aptitude. What this common instructional approach can potentially do, therefore, is address the portion of the initial skill gap that is related to the former and not the latter, so that what is simply a difference in preschool opportunity does not calcify into long-term school disparities.

The Present Study

The starting point for this study was the well-documented pattern of initial differences in skill levels at the beginning of elementary school developing into divergent trajectories of achievement. Our expectation was that this divergence would be less pronounced when children had positive relations with elementary school teachers leading classrooms in high-order skill-learning activities. In other words, children who started school behind their peers would catch up over time when they were in classrooms emphasizing higher order skill development and had supportive relations with the teachers in these classrooms. They would fall further behind when in classrooms in which instruction focused solely on basic skill development or in which higher order skill development was not coupled with positive teacher–student relations.

Method

Data and Sample

The SECCYD is a comprehensive longitudinal study that was originally designed to answer questions about the relation between child care and child development (see http://secc.rti.org for more details) but has evolved into a study of the general adjustment and functioning of young people. Data collection began in 1991 in 10 locations—Little Rock, AR; Irvine, CA; Lawrence, KS; Boston, MA; Philadelphia, PA; Pittsburgh, PA; Charlottesville, VA; Morganton, NC; Seattle, WA; and Madison, WI. During selected 24-hr sampling periods, study personnel visited new mothers in the hospital in which they had just given birth. To be eligible, the mother had to be over 18 years of age, healthy, and conversant in English, and the infant had to be a singleton. A month after the birth of the target infant, 1,364 families (58% of those contacted) were enrolled in the study. The collection of data from parents, children, and other adults (e.g., child-care providers, teachers) in home, laboratory playroom, and other visits proceeded in several stages from birth and is ongoing.

This study used an SECCYD subsample of all children who had their classrooms assessed in either third or fifth grade (n = 587). These children were spread across 474 classrooms in 471 schools in third grade, with a maximum of two children in 38 classrooms (the corresponding numbers for fifth grade: 430, 427, 54, respectively). Thus, although about 20% of the children were technically clustered within classrooms and schools in either third or fifth grade, they were not clustered at a rate that complicated statistical inference by, for example, deflating standard errors through large violations of assumptions of independence (Guilkey & Murphy, 1993).

Our focus on a subsample of the data set introduces the question of selection bias. In other words, did the cases in the analytical sample differ systematically from those in the original SECCYD sample who left the study over time or did not have their classrooms observed repeatedly? Could any such differences bias results? Past SECCYD research has indicated that sample bias related to panel attrition and instrument-specific missingness is not substantial (NICHD Early Child Care Research Network, 2005a). Our own descriptive analysis (reported shortly) confirms this general conclusion.

Measures

Descriptive statistics for all study variables are presented in the first column of Table 1.

Academic achievement and early skill groups—To measure academic achievement, we focused exclusively on test scores in math. The use of a standardized test facilitated comparison across the hundreds of schools attended by children in the SECCYD sample. The focus on math (rather than, e.g., language) was motivated by the lower overall exposure of children to math enrichment prior to elementary school and by the highly cumulative nature of math instruction in elementary school (Ginsburg, 1989).

The children in the sample took the Applied Problems test of the Woodcock–Johnson Psychoeducational Battery–Revised (WJ-R; Woodcock & Johnson, 1989) in lab visits at the 54-month data collection and then again in various school grades. This test includes math calculations, including calculations in response to word and story problems that require both basic and conceptual skills. Raw scores were converted to *W* scores, which are special transformations of the Rasch ability scale that center the raw score on a value of 500 to ease comparisons across standardized tests. The average test score in our sample rose from just under 500 in third grade to just over 500 in fifth grade (a spread of about 17 points), representing the expected learning gains that come with additional years of math instruction.

Recall that the major focus of this study was on children who differed in math skills at school entry. Because of our expectation that low-skilled children would show differential benefits of classroom factors (vs. the rest of the entry-level skill distribution), we decided to use a categorization scheme rather than using the continuous measure of entry-level skills. In addition to the conceptual rationale, a categorization scheme also helps findings be more policy relevant (Kraemer, Stice, Kazdin, Offord, & Kupfer, 2001). The 54-month WJ-R scores were the basis of this scheme. After testing several approaches, we settled on a standard deviation unit system that created distinct groups and also aligned with service provision recommendations in schools. Children who scored one standard deviation below the population mean on the Applied Problems test (standard score lower than 85) were designated as the low early skill group (*n* = 71), those who scored between one standard deviation below and one standard deviation above the population mean (standard score of 85 to115) were designated as average the early skill group (*n* = 396), and those who scored one standard deviation above the population mean or higher (standard score greater than 115) were designated as the high early skill group (*n* = 120). Scores on subsequent tests were then used to estimate growth curves of math achievement during elementary school (see plan of analysis described later) for each of these three early skill groups.

The distributions of the test scores serving as the outcomes were normal. Before describing the numerous variables used to predict the outcome, we should note that all had approximately symmetric distributions and that none had a highly skewed or flat distribution.

Classroom instruction—The Classroom Observation System (COS) was developed to provide multiple measures of the quality of elementary school classrooms (see Hamre & Pianta, 2005; Pianta, Belsky, Houts, Morrison, & the NICHD Early Child Care Research Network, 2007). COS scales are classroom-level measures and are not specific to any one child in the classroom. In the first-, third-, and fifth-grade data collections, classroom observations were recorded in 60 blocks (30 s to observe, 30 s to record) over two cycles. Observations were made by centrally trained and certified personnel, and extensive testing was conducted to ensure an overall reliability rating of nearly .80 (correlation with master coders' ratings).

Among the many COS scales are measures of the frequency of inferential instructional techniques and of the frequency of basic skill instructional techniques in the classroom.[2] What we refer to as inferential or higher order instruction captures a broad dimension of academic activities encompassing analysis, inference, and synthesis. Observations of classroom activities that required students to reason, solve problems, create, evaluate, or engage in deductive reasoning were coded as such. For example, a teacher might have asked students to explain how they arrived at some math solution or encouraged students to use their own errors on a math story problem to cultivate a better understanding of some math concept. Alternatively, when classroom activities required students to come up with yes–no responses or responses that were either correct or incorrect, an observation of basic skills instruction was recorded. For example, a teacher might have done math drills with students or drawn shapes on the board and asked students to name them.

Observers noted the frequency of each kind of teacher instructional behavior across cycles, so that the final scores for each represent the count of the relevant behavior observed but not the duration. The mean for basic skills instruction in Table 1 indicates that personnel observed an instance of basic skills instruction, on average, in just over 20 of the 60 time

[2]The inference and basic skill distinctions could not be made in first grade. Therefore, we used only the third- and fifth-grade measures in analyses.

blocks during their third-grade observation period (over 18 for fifth grade). Clearly, this classroom activity far exceeded the rate of inference-based instruction during the same observation periods (noted in fewer than 3 of the 60 blocks in third grade and fewer than 4 in fifth grade).

Teacher–student relations—The classroom teachers completed a 15-item questionnaire to assess teacher–student emotional closeness ($\alpha = .85$) and conflict ($\alpha = .91$). Examples of the eight items in the Closeness scale are "I share an affectionate, warm relationship with this child" and "When I praise this child, he/she beams with pride." Likert scores—ranging from 1 (*definitely does not apply*) to 5 (*definitely applies*)—were summed across the items (M = 33.13 for third grade and 31.96 for fifth grade in Table 1). Examples of the seven items in the Conflict scale are "This child and I always seem to struggle with each other" and "Dealing with this child drains my energy." The same summation procedure as the Closeness scale was followed (M = 11.86 and 11.41, respectively). Unlike the COS scales, the teacher–student variables are specific to individual SECCYD children and not to their classrooms. They tap an actual relationship rather than classroom climate.

Demographic controls—Several characteristics were measured to control for demographic variability in math skills and classroom placement: race (1 = *White*, 0 = *non-White*), age (in years), and years of maternal education. Given scholarly and public interest in gender differences in math achievement (Eccles, 1994), we also controlled for a binary marker of gender (1 = *female*, 0 = *male*). According to Table 1, the sample was primarily White (78%) with a mean level of maternal education between high school and college graduation. The children in the sample were split nearly evenly by gender (slightly favoring girls) and were, on average, nearly 9 years old. Another important control was for study site to account for any nonrandom differences across the 10 data collection sites. Dummy variables for the sites were included in all analyses, with the largest site serving as the reference.

School controls—The National Center for Education Statistics database on American schools allowed the measurement of two demographic aspects of schools attended by SECCYD children: proportion of the student body receiving free or reduced price lunch (a proxy for school poverty) and non-White proportion of student body. These two school characteristics were combined to form a measure of school demographic risk because they were highly correlated ($r = .75$). Reflecting broad patterns of school segregation in the United States, students in the sample—who were mostly White and middle class—attended schools with low rates of minority/poverty enrollment. This school-level control accounts for the different school contexts in which the focal classrooms were situated.

Plan of Analysis

Before laying out the analytical format of this study, we need to address the possible selection bias—mentioned earlier— introduced by longitudinal sample attrition and the nonnegligible missing data on classroom measures. According to the statistics in Table 2, the SECCYD children included in our analytical sample differed from those excluded from it (primarily for not having their classrooms assessed multiple times) on only two measures. The analytical sample children were slightly younger when they took the WJ-R tests and attended schools with a slightly higher percentage of non-White students. They did not, however, differ from the other children on the focal study variables, including math test scores, classroom instruction, and teacher–student relations. We performed one other methodological check on potential sampling bias, which we describe shortly.

After some preliminary analyses (e.g., descriptive statistics by early skill group, correlations among focal variables), we used the mixed procedure in SAS to conduct a repeated measures analysis with random intercept hierarchical linear modeling (HLM). Two random-intercept models were fit. These models included repeated measures on the outcome (math test scores in third grade and fifth grade),[3] classroom instruction, and teacher–student relations and accounted for the dependency in these repeated measures by estimating a random intercept for each child.[4] The child's age at each assessment was used to index the repeated factor and was included in all analyses. The children's 54-month skill level group was the predictor of interest, and the Age × Skill Group interaction gauged whether rates of change over time varied across these three groups. These analyses proceeded in two steps.

First, we were interested in determining whether test score changes over time were different for children in the three early skill groups (e.g., the Age × Skill Group interaction). That is, the first model—with age, 54-month skill group, Age × Skill Group, and time-invariant demographic covariates—described the extent to which children exhibited different trajectories of math achievement if they entered school with different 54-month skill levels. Also worth noting is that the demographic covariates were not crossed with skill group, meaning that their regression parameters and standard errors were equivalent across groups in the final model. This model tested the main effects of age and early skill group as well as interactions between the two, which gave estimates of the intercepts (overall level) and slopes (rates of change) in math achievement over time for each early skill group. The significance levels of observed differences in these intercepts and slopes were assessed with *F* tests.

Second, we were interested in whether the associations of classroom instruction and teacher–student relations with math achievement over time varied according to entry-level skills in math. That is, the second HLM included the school, classroom instruction, and teacher–student relations factors as predictors. Each of these three sets of variables was allowed to interact with age and early skill group. We also added interactions between the two classroom instructional variables and the two teacher–student relations variables to determine whether the match between instructional elements of the classroom and teacher–student relations predicted math achievement over time differently across early skill groups (see the Appendix for equations). Again, significant differences across groups were assessed with *F* tests.

To return to the sampling bias issue, we did a post hoc test akin to a Heckman correction. This test took the form of a two-stage regression. In the first stage, logistic regressions estimated the likelihood that a case in the full SECCYD sample was included in our analytical sample as a function of a large number of predictors, including all used in our main models. The predicted probability was then entered into our main models as a propensity weight. We could then interpret the results of this modeling approach as the effects of focal predictors (e.g., classroom instructional factors, teacher–student relations factors, early skill groups) on the outcome, controlling for the likelihood of a student being included in the analytical sample. For these focal predictors, the results of this modeling strategy produced the same results as the earlier models. The only change was that the coefficients for a few of the nonfocal covariates switched from being almost significant to being significant at conventional thresholds.

[3]Although achievement was also tested in first grade, we could not include these data in the models because the two focal classroom variables were available only in third and fifth grade.

[4]Classroom- and school-level estimates of variance were not modeled given the small number of children clustered in the same classrooms and schools. This was a matter of the nonclustered nature of the data, not the level on which these variables were measured. They were unique to schools and classrooms.

Results

Comparison of Early Skill Groups

The three categories of early math skill (low, average, high) were defined by children's scores on the WJ-R Applied Problems test prior to their entry into elementary school (at 54 months of age). Table 1 provides a comparison of these three groups on key variables. Beginning with demographic characteristics, children in the low early skill group were less likely than other children to be White, and they had mothers with lower levels of educational attainment. Turning to school, classroom, and teacher–student factors, children in the low early skill group received more basic skills instruction in their third-and fifth-grade classrooms than children in the other two skill groups. They also tended to be less close to their teachers and to experience more conflict with them. Finally, once in school, these low early skill children scored below their peers on math tests each year. In contrast, children with the highest level of skills when they entered elementary school had the most educated mothers, received less basic skill instruction, and had less conflicted relations with their teachers. They always posted the highest math test scores as they moved through elementary school.

Correlations provide another preliminary picture of the data (Table 3). Math scores in third and fifth grade were higher when children were White (*r*s = .26 in third grade, .29 in fifth grade), had more educated mothers (*r*s = .37 in third, .41 in fifth), and attended schools with fewer poor and race/ethnic minority students (*r*s = −.22 in third, −.33 in fifth). Children scored better on third-grade math tests when they had warmer relations with teachers in third grade (*r* = .16), when they had less conflict with teachers in both grades (*r*s = −.35 in third, −.26 in fifth), and when their teachers spent less time on basic skills instruction (*r* = −.15) in fifth grade. Also worth noting is the moderate-to-modest negative associations between the two teacher–student measures (*r*s = −.41 in third, −.35 in fifth) and between the two classroom instruction measures (*r*s = −.20 in third, −.10 in fifth). Finally, the two teacher–student relations measures were not significantly correlated with the two classroom instruction measures.

In summary, children who had lower math skills entering elementary school were disproportionately of low socioeconomic status and/or race/ethnic minority status. They tended to go into classrooms in which basic skills instruction was the norm, and they had more difficult relations with teachers, which put them at academic risk. This observed pattern needs to be examined in a multivariate framework, and the potential combined role of instructional classroom characteristics and socioemotional teacher–student relations (not just their independent roles) needs to be considered. In other words, the next step was to examine the interactive role of classroom instruction and teacher–student relations in predicting math achievement over time while controlling for important confounds.

Early Skills, Classrooms, and Math Achievement Gains

To pursue this multivariate strategy, we estimated a series of repeated measures models for math achievement (see Table 4). Model 1 had demographic characteristics as time-invariant covariates, with main effects and interactions between age and early skill group (measured at 54 months).[5] With the exception of the Early Skill Group × Age interaction, all of these variables significantly predicted the outcome in expected ways.

[5]Variables were centered at the grand mean for the stacked third-grade/fifth-grade sample so that main effects could be interpreted within the range of the observed data.

The association between early skill group and math achievement over time is the most important part of Model 1 to consider. To provide a clearer picture of what the coefficients in Model 1 represent, we present the predicted regression lines for each of the three early skill groups in Figure 1. These lines were based on the estimated intercept and slope for math achievement within each early skill group from the HLM analysis, net of the control variables. Beginning with math achievement level at the average age, children in the low early skill group scored lower on math tests than those in the average early skill group (d = 1.11) and those in the high early skill group (d = 1.59).[6] At the same time, children in the average early skill group scored lower than children in the high early skill group (d = 0.48). Turning to rates of change, the slope of the trajectories between third and fifth grades did not differ significantly among the three early skill groups.

In summary, this initial modeling step indicated that children who entered elementary school with different skill levels in math followed different but parallel trajectories of math achievement through fifth grade. They started at different levels and maintained that difference—without adding to it—over time.

Model 2 in Table 4 added the school, classroom instruction, and teacher–student factors to Model 1. As discussed in the plan of analysis, we crossed each of these variables with both age and early skill group and then also crossed the teacher–student relations and instructional styles. Results indicated that children in the three early skill groups continued to differ substantially in math achievement. The magnitude of these differences, however, was smaller (according to the F test), most likely because of the increased variability within skill groups related to significant interaction terms.

Overall, children in the low early skill group scored lower on math tests over time than their peers in the average early skill group (d = 1.02) and in the high early skill group (d = 1.51), with the high early skill group also outscoring the average early skill group (d = 0.49). Teacher–student closeness was positively associated with math achievement in third and fifth grades but more so in third grade than fifth grade (Age × Teacher–Student Closeness: $B = -.11$, $p < .05$). The three-way interaction among teacher–student conflict, inference instruction, and early skill group was statistically significant ($F = 3.20$, $p < .05$, with $B = .10$, $p < .05$, for low early skill group).[7]

This interaction was explored by graphing predicted math achievement levels based on the HLM parameters. Using one standard deviation above the mean to represent high scores on the math test and one standard deviation below the mean to represent low scores on this test, we estimated regression lines for children in each early skill group whose classrooms were characterized by either high or low inferential skill instruction and who had either high or low conflict with their teachers in these classrooms (with all other covariates and predictors held to their sample means). Doing so revealed that low levels of inferential instruction in the classroom were related to significantly lower test scores for children in the low entry skill group if they also experienced high levels of teacher–child conflict. To be more specific, inference skills instruction and teacher–student conflict predicted math achievement more strongly for the low early skill group than for the other two groups. The children in the low early skill group who experienced relatively frequent inference skills

[6]The Cohen statistic for comparing adjusted means (intercepts here): (Mean 1 − Mean 2)/standard deviation. The standard deviation was the estimated variability for the individual intercept.

[7]These analyses were re-estimated with the grand-mean centered 54-month WJ-R score and the squared score replacing the early skill categorization. Results suggested that entry-level math skills had linear and quadratic associations with math skills in third or fifth grade and that teacher closeness moderated the association between entry-level math skills and math skills at third or fifth grade, such that teacher closeness was a stronger predictor when entry-level math skills were lower than when they were higher. In addition, those analyses indicate main effects for conflict with the teacher and interactions among age and teacher closeness and teacher conflict.

instruction showed more rapid gains in math achievement than children in the low early skill group who experienced relatively infrequent high inference instruction. This association between high inference skills instruction and rate of math achievement change over time was greater for children in the low early skill groups than for children in the other two groups.

Figure 2 presents these results for different subgroups of low early skill children. Within this group, children in classrooms characterized by high levels of inference skills instruction who had low levels of conflict with teachers in these classrooms demonstrated higher absolute levels of math achievement in elementary school and posted higher rates of increase in math achievement over time than their low early skill group peers in classrooms with low levels of inference skills instruction who had low conflict with teachers. They also did better over time than their low early skill group peers in classrooms characterized by high levels of inference skills instruction who had high conflict with teachers. This focal group—low early skill children in classrooms with higher order instruction who had better teacher–student relations—had trajectories of math achievement during elementary school with a more pronounced upward slope than children in the average and high early skill groups (not shown in Figure 2). In other words, these children who looked relatively unskilled at the start of elementary school made more math progress as elementary school unfolded than their equally unskilled peers who did not receive these educational benefits.

Discussion and Conclusion

In the SECCYD, children who entered the educational system with already well-developed math skills maintained their advantages over their peers who had less preparation in the preschool period. Their less prepared peers did not catch up. Of importance, this individual-level pattern is also a population-level pattern. Because the children in the high early skill group were disproportionately from White families with high socioeconomic status and their peers with low early skills were disproportionately from non-White families with low socioeconomic status, this cumulative process helped to reinforce existing differences between advantaged and disadvantaged segments of the U.S. population.

Such patterns of advantage and disadvantage have long been the focus of a good deal of social policy—most recently, the No Child Left Behind Act of 2001—and of scientific research aiming to inform that policy (Gamoran, 2007). Clearly, the most direct remedy is to target the extreme disparities in school readiness that fuel persistent disadvantage, disparities that are, in turn, rooted in unequal access to opportunities to learn in the major contexts of early childhood. Research demonstrating the leveling effect of high-quality child care and early enrichment programs certainly suggests the potential payoff of this approach; indeed, actual programs (e.g., Head Start, universal pre-kindergarten) have been enacted and are translating this potential into reality (Clarke-Stewart & Allhusen, 2005; Gormley, Gayer, Phillips, & Dawson, 2005; Keating & Simonton, 2008; Ludwig & Miller, 2007; Zigler & Styfco, 1993). Yet, what occurs prior to school is only one part of the equation. What happens after school entry also matters.

After school entry, the cycle of persistent advantage and disadvantage is maintained in the classroom, and so it is also likely that the classroom is the place where the cycle can be broken. The results of this study provide some evidence for what has long been suggested, namely, that moving from a more exclusively basic skills pedagogy in elementary school classrooms to a higher order system of instruction that incorporates skills into conceptual pedagogy can benefit students at the low end of the preparedness/ability spectrum just as much as those at the high end (Connor et al., 2004; Griffin, 2004). This finding echoes evidence from the U.S. Catholic school literature and the detracking literature (Bryk et al., 1993; Cohen, 1994). Rather than channeling children who come to school with less preparation—who often come from the most disadvantaged segments of the population—into classes that sacrifice challenging lesson plans in even a well-intentioned attempt to get their basic skills up to speed, exposing these children to some of the activities that their peers receive can actually enhance their learning in ways that allow them to make up ground. In the latter situation, these children appear to achieve their basic skills and then some.

Yet, this approach to helping less school-ready children catch up is likely quite delicate. Relations between the teacher and individual students in the classroom factor into how much students gain from classroom instruction. More specifically, teacher–student conflict seemed to be an important dimension to consider when looking at possible learning gains related to higher order instruction among children with low levels of early skills. A student who comes into a classroom with less academic preparation or lower aptitude might experience frustration or distress when pushed to develop higher order skills along with his or her classmates. Similarly, a teacher trying to implement a strategy for developing such skills in the classroom might become exasperated with a student who is struggling. These feelings—especially when they go both ways—likely reduce the potential for higher order instruction to pull up the achievement of students with less developed skills. Thus, the key to this approach appears to be balance between teaching subject and teaching style. Programs that have proven success in facilitating this balance—for example, complex instruction techniques implemented in detracked classrooms (Cohen, 1994)—should then be prioritized in the early grades, especially in classrooms with mixed skill levels. Teachers charged with implementing higher order instructional strategies in mixed-skill classrooms should also receive training in how to manage conflict with students, and evaluations should take into account both the use of such instructional practices on the classroom level and the interpersonal tone of instruction and interaction on the level of individual students.

Of course, this basic conclusion comes with several caveats. First, the leveling effects of classroom instruction and teacher–student relations, whether combined or separate, were not large in magnitude. In other words, the low early skill children experiencing the most favorable educational combination made up only a modest portion of the achievement gap compared with their high early skill peers. They certainly did not close the gap, which is not surprising given that achievement gaps are both robust and multidetermined. The gap, on average, is both substantial in magnitude and durable (as shown in Figure 1), indicating the need for multifaceted policy approaches. Second, this conclusion concerns only math, not other core domains. Clearly, researchers need to investigate whether this approach might be useful in general and not just in relation to one subject (although a very important subject). Third, worth noting is that the WJ-R test tapped math problem solving. To the extent that such problem solving drew fairly evenly on basic and inferential/analytical skills, this test may be particularly sensitive to the kinds of inferential and deductive strategies tapped by our classroom high-order instruction scale. To the extent that such problem solving was primarily an exercise in basic skills, however, the test might be insufficiently sensitive to the kinds of instruction that we were targeting. Thus, looking at how these scales are associated with multiple forms of achievement (e.g., grades, engagement) in multiple subjects will provide a more comprehensive test of our conceptual model. Fourth, these findings are based on observational data. Although our longitudinal design and control for numerous confounding factors certainly improves causal inference, we certainly cannot establish causality here. This study, therefore, should be seen as an attempt to inform future experimental investigations.

Even with these caveats in mind, however, the findings of this study are informative. They reiterate the value of integrating the two sides of schooling—the formal, instructional, curricular aspects of school and the informal, social, interactional aspects of school (Coleman, 1990; Greenberg et al., 2003). Focusing on the former without respecting the power of the latter can take efforts to educate children and solve systemic inequities only so far. This message is a timely one in the No Child Left Behind era in which schools are held to increasingly high standards of performance.

Acknowledgments

The NICHD Study of Early Child Care and Youth Development is directed by a steering committee and supported by NICHD through a cooperative agreement (U10), which calls for scientific collaboration between the grantees and the NICHD staff. The content is solely the responsibility of the named authors and does not necessarily represent the official views of the NICHD, the National Institutes of Health, or individual members of the network. Current members of the Steering Committee of the NICHD Early Child Care Research Network, listed in alphabetical order, are Jay Belsky (Birkbeck University of London), Cathryn Booth-LaForce (University of Washington), Robert H. Bradley (University of Arkansas at Little Rock), Celia A. Brownell (University of Pittsburgh), Margaret Burchinal (University of North Carolina at Chapel Hill), Susan B. Campbell (University of Pittsburgh), Elizabeth Cauffman (University of California, Irvine), Alison Clarke-Stewart (University of California, Irvine), Martha Cox (University of North Carolina at Chapel Hill), Robert Crosnoe (University of Texas at Austin), James A. Griffin (NICHD project scientist and scientific coordinator), Bonnie Halpern-Felsher (University of California, San Francisco), Willard Hartup (University of Minnesota), Kathryn Hirsh-Pasek (Temple University), Daniel Keating (University of Michigan), Bonnie Knoke (RTI International), Tama Leventhal (Tufts University), Kathleen McCartney (Harvard University), Vonnie C. McLoyd (University of North Carolina at Chapel Hill), Fred Morrison (University of Michigan), Philip Nader (University of California, San Diego), Marion O'Brien (University of North Carolina at Greensboro), Margaret Tresch Owen (University of Texas at Dallas), Ross Parke (University of California, Riverside), Robert Pianta (University of Virginia), Kim M. Pierce (University of Wisconsin—Madison), A. Vijaya Rao (RTI International), Glenn I. Roisman (University of Illinois at Urbana-Champaign), Susan Spieker (University of Washington), Laurence Steinberg (Temple University), Elizabeth Susman (Pennsylvania State University), Deborah Lowe Vandell (University of California, Irvine), and Marsha Weinraub (Temple University).

References

Alexander KL, Entwisle DR. Achievement in the first two years of school: Patterns and processes. Monographs of the Society for Research in Child Development 1988;53:1–157. [PubMed: 3226426]

Bryk, AS.; Lee, VE.; Holland, PB. Catholic schools and the common good. Cambridge, MA: Harvard Press; 1993.

Cameron CE, Connor CM, Morrison FJ. Effects of variation in teacher organization on classroom functioning. Journal of School Psychology 2005;43:61–85.

Ceci SJ, Papierno PB. The rhetoric and reality of gap closing: When the "have nots" gain but the "haves" gain even more. American Psychologist 2005;60:149–160. [PubMed: 15740447]

Clarke-Stewart, A.; Allhusen, VD. What we know about childcare. Cambridge, MA: Harvard University Press; 2005.

Cohen, EG. Designing groupwork: Strategies for the heterogeneous classroom. New York, NY: Teachers College; 1994.

Coleman, JS. Foundations of social theory. Cambridge, MA: Harvard; 1990.

Connor CM, Morrison FJ, Fishman BJ, Schatschneider C, Underwood P. The early years: Algorithm-guided individualized reading instruction. Science 2007 January 26;315:464–465. [PubMed: 17255498]

Connor CM, Morrison FJ, Katch LE. Beyond the reading wars: Exploring the effects of child-instruction interactions on growth in early reading. Scientific Studies of Reading 2004;8:305–336.

Crosnoe, R. Mexican roots, American schools: Helping Mexican immigrant children succeed. Palo Alto, CA: Stanford University Press; 2006.

Eccles JS. Understanding women's educational and occupational choices: Applying Eccles et al. model of achievement-related choices. Psychology of Women Quarterly 1994;18:585–610.

Entwisle, DR.; Alexander, KL. The first grade transition in life course perspective. In: Mortimer, J.; Shanahan, M., editors. Handbook of the life course. New York, NY: Kluwer Academic/Plenum; 2002. p. 229-250.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Eunice Kennedy Shriver National Institute of Child Health and Human Development Early Child Care Research Network. Child care and child development. New York, NY: Guilford Press; 2005a.

Eunice Kennedy Shriver National Institute of Child Health and Human Development Early Child Care Research Network. Duration and developmental timing of poverty and children's cognitive and social development from birth through third grade; Child Development. 2005b. p. 795-810.

Farkas, G. Human capital or cultural capital? Ethnicity and poverty groups in an urban school district. New York, NY: de Gruyter; 1996.

Gamoran, A., editor. Standards-based reform and the poverty gap: Lessons for No Child Left Behind. Washington, DC: Brookings Institution; 2007.

Ginsburg, HP. Children's arithmetic: How they learn it and how you teach it. Austin, TX: Pro Ed.; 1989.

Gormley W, Gayer T, Phillips D, Dawson B. The effects of universal pre-K on cognitive development. Developmental Psychology 2005;41:872–884. [PubMed: 16351334]

Greenberg MT, Weissberg RP, O'Brien MU, Zins JE, Fredericks L, Resnik H, Elias MJ. Enhancing school-based prevention and youth development through coordinated social, emotional, and academic learning. American Psychologist 2003;58:466–474. [PubMed: 12971193]

Griffin S. Building number sense with Number Worlds: A mathematics program for young children. Early Childhood Research Quarterly 2004;19:173–180.

Guilkey D, Murphy J. Estimation and testing in the random effects probit model. Journal of Econometrics 1993;59:301–317.

Hamre B, Pianta RC. Can instructional and emotional support in the first grade classroom make a difference for children at risk of school failure? Child Development 2005;76:949–967. [PubMed: 16149994]

Howes C. Social–emotional classroom climate in child care, child–teacher relationships, and children's second grade peer relations. Social Development 2000;9:191–204.

Keating, DP.; Simonton, SZ. Developmental health effects of human development policies. In: Schoeni, R.; House, J.; Kaplan, G.; Pollack, H., editors. Social and economic policy as health policy. New York, NY: Sage; 2008. p. 61-94.

Kraemer HC, Stice E, Kazdin A, Offord D, Kupfer D. How do risk factors work together? Mediators, moderators, and independent, overlapping, and proxy risk factors. American Journal of Psychiatry 2001;158:848–856. [PubMed: 11384888]

Lee, V.; Burkham, D. Inequality at the starting gate: Social background differences in achievement as children begin school. Washington DC: Economic Policy Institute; 2002.

Ludwig J, Miller DL. Does Head Start improve children's life chances? Evidence from a regression–discontinuity design. Quarterly Journal of Economics 2007;122:159–208.

Magnuson KA, Meyers MK, Ruhm CJ, Waldfogel J. Inequality in preschool education and school readiness. American Educational Research Journal 2004;41:115–158.

No Child Left Behind Act of 2001. 2002 Pub. L. No. 107-110, 115 Stat 1425.

Pianta RC, Belsky J, Houts R, Morrison F, Clifford R, Early D, Barbarin O. Features of pre-kindergarten programs, classrooms, and teachers: Do they predict observed classroom quality and child–teacher interactions. Applied Developmental Science 2005;9:144–159.

Pianta RC, Belsky J, Houts R, Morrison F. the NICHD Early Child Care Research Network. Opportunities to learn in America's elementary classrooms. Science 2007 March 30;315:1795–1796. [PubMed: 17395814]

Pianta, RC.; Cox, MJ. The transition to kindergarten. Baltimore, MD: Brookes; 1999.

Pianta RC, Steinberg MS, Rollins KB. The first two years of school: Teacher–child relationships and deflections in children's classroom adjustment. Development and Psychopathology 1995;7:295–312.

Raver, CC.; Gershoff, E.; Aber, L. Child Development. Vol. 78. 2007. Testing equivalence of mediating models of income, parenting, and school readiness for White, Black, and Hispanic children in a national sample; p. 96-115.

Rimm-Kaufman S, La Paro K, Pianta RC, Downer J. The contribution of classroom setting and quality of instruction to children's behavior in kindergarten classrooms. Elementary School Journal 2005;105:377–394.

Sadowski M. The school readiness gap. Harvard Education Letter 2006 July/August;22(4):1–2.

Waldfogel, J. What children need. Cambridge, MA: Harvard University Press; 2006.

Woodcock, RW.; Johnson, MB. Woodcock–Johnson Revised Tests of Achievement. Itasca, IL: Riverside; 1989.

Xue Y, Meisels SJ. Early literacy instruction and learning in kindergarten: Evidence from the Early Childhood Longitudinal Study–Kindergarten Class of 1998–1999. American Educational Research Journal 2004;41:191–229.

Zigler EF, Styfco SJ. Using research and theory to justify and inform Head Start expansion. SRCD Social Policy Report 1993;7:1–24.

Appendix

HLM Modeling Steps

The following equation shows the HLM modeling steps. In this equation, i = ith individual, and j = jth occasion. Note that interactions among classroom and teacher characteristics are included in the vector of classroom and teacher characteristics.

$$Y_{ij}=\gamma_{i0}+\gamma_{i1}\text{ age}+\gamma_{2}\text{ time}-\text{varying covariates}_{ij}+\varepsilon_{ij}$$
$$\gamma_{i0}=B_{10}+B_{11}\text{ early skill group}_{i}+B_{12}\text{ classroom and teacher characteristics}_{ij}+B_{13}\text{ early skills group }\times\text{ classroom and teacher characteristics}_{ij}+e_{i}$$
$$\gamma_{i1}=B_{20}+B_{21}\text{ early skill group}_{i}+B_{22}\text{ classroom and teacher characteristics}_{ij}+B_{23}\text{ early skills group }\times\text{ classroom and teacher characteristics}_{ij}$$

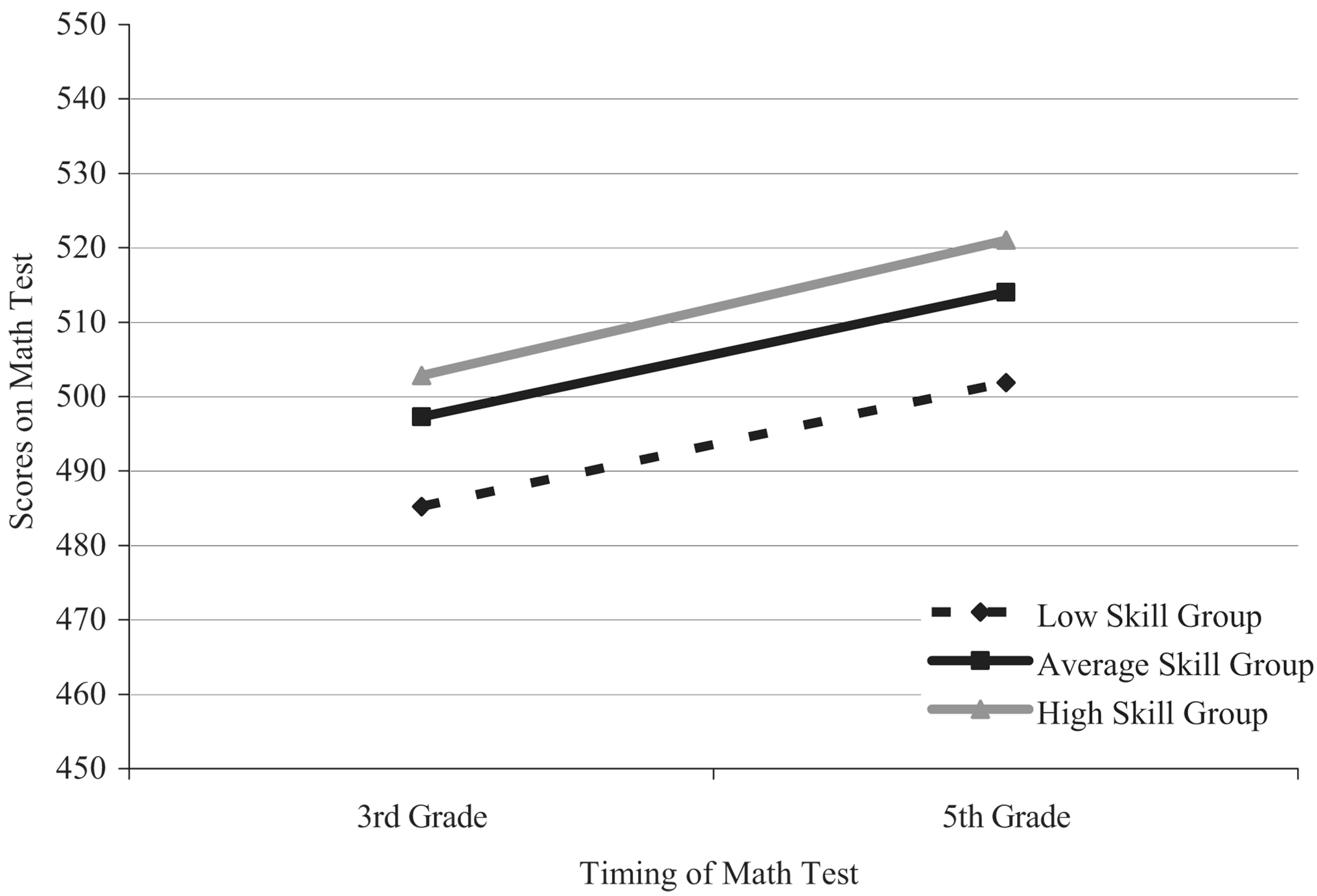

Figure 1.
Math test scores in elementary school in three 54-month skill groups.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

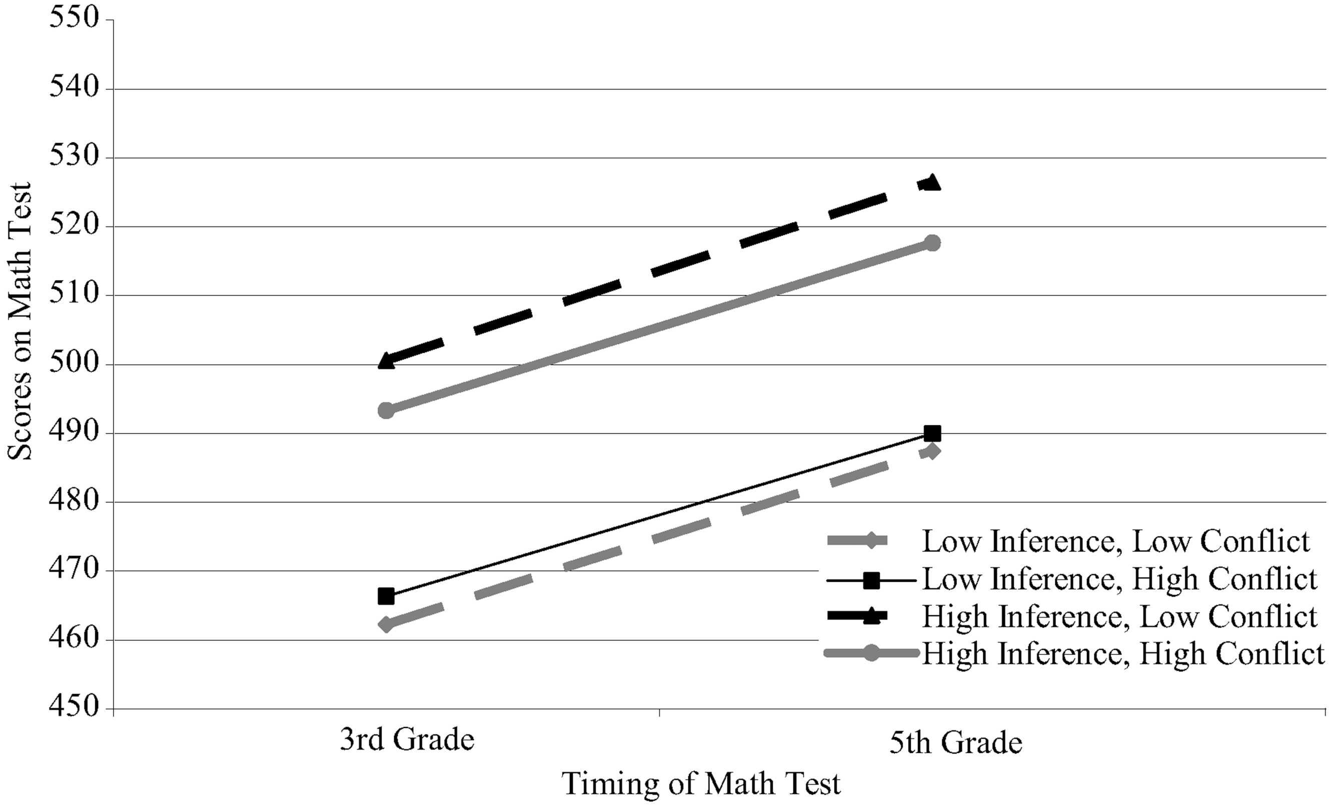

Figure 2.

Math test score by classroom inference instruction and teacher–student conflict in the 54-month low-skill group.

Table 1

Descriptive Statistics for Focal Study Variables, by Early Skill Group

Focal study variable	M (SD) for full sample (n = 587)	M (SD) by 54-month skill group: Low[a] (n = 71)	Average[b] (n = 396)	High[c] (n = 120)	F(2, 585)
Demographic characteristics					
Race (proportion White)	0.78	0.46[c]	0.80[b]	0.94[a]	61.9*** [d]
Gender (proportion male)	0.49	0.59	0.47	0.43	4.63[d]
Age	8.98 (0.30)	9.02 (0.36)	8.96 (0.29)	8.94 (0.30)	1.77
Maternal education	14.16 (2.37)	12.62[c] (1.99)	13.95[b] (2.26)	15.70[a] (2.06)	49.1***
Third-grade classroom/teacher factors					
Basic skills instruction	20.16 (9.62)	22.17[b] (10.24)	20.33[a,b](9.56)	18.54[a] (9.64)	3.28*
Inference skills instruction	2.42 (3.81)	2.26 (3.93)	2.24 (3.63)	3.19 (4.36)	2.95
Teacher–student closeness	33.13 (5.01)	31.20[b] (5.88)	33.24[a] (4.85)	33.89[a] (4.63)	6.94**
Teacher–student conflict	11.86 (6.30)	15.80[a] (8.59)	11.57[b] (6.00)	10.02[c] (4.28)	20.85***
Third-grade school characteristics					
Third-grade school demographic risk	0.27 (0.22)	0.41[a] (0.26)	0.27[b] (0.21)	0.19[c] (0.17)	25.5***
Fifth-grade school demographic risk	0.31 (0.23)	0.45[a] (0.27)	0.31[b] (0.22)	0.23[c] (0.19)	21.9***[e]
Math achievement					
Third-grade math test score	493.0 (13.13)	479.0[c] (17.3)	493.8[b] (10.7)	500.0[a] (8.7)	77.47***
Fifth-grade math test score	510.2 (14.13)	495.0[c] (16.0)	511.2[b] (10.9)	518.4[a] (9.0)	88.79***[e]

Note. Means with different subscripts differed significantly according to F tests.

[a] The low early skill group scored less than one standard deviation below the mean on the 54-month Woodcock–Johnson Revised Applied Problems test.

[b] The average early skill group scored within one standard deviation of the mean.

[c] The high early skill group scored more than one standard deviation above the mean.

[d] Chi-square test reported (df = 1, n = 588).

[e] The df-e for fifth-grade measures was 515.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

Table 2

Comparison of Analytical Sample With Attrition Sample

	M **(*SD*)**		
Continuous variable	**Analysis sample (*n* = 587)**	**Attrition sample (*n* = 777)**	**Test of difference (*t* test)**
Maternal education	14.1 (2.4)	14.3 (2.6)	*ns*
WJ-R math test score	493.3 (12.8)	493.8 (12.8)	*ns*
Age at WJ-R math test	8.97 (0.30)	9.02 (0.32)	−2.78**
Percentage school free/reduced lunch	0.29 (0.23)	0.30 (0.24)	*ns*
Percentage of school non-White	0.25 (0.24)	0.17 (0.25)	5.42***
Basic skills instruction	20.2 (9.7)	19.4 (8.8)	*ns*
Inference skills instruction	2.43 (3.84)	2.18 (3.72)	*ns*
Teacher–student closeness	33.1 (5.0)	33.0 (5.4)	*ns*
Teacher–student conflict	11.8 (6.3)	11.4 (5.6)	*ns*

Note. Chi-square tests were conducted for comparisons of two binary variables. The proportion of White students did not differ significantly across the two samples (79% for analytical, 75% for attrition), but the proportion of male students was significantly lower in the analysis sample (48%) than the attrition sample (54%).

WJ-R = Woodcock–Johnson Psychoeducational Battery–Revised

** $p < .01$.

*** $p < .001$.

Table 3

Correlations Among Focal Study Variables (n = 587)

Variable	1	2	3	4	5	6	7	8	9	10
1. Math test score	—	.03	.37***	.26***	.04	−.22***	−.15***	.04	.16***	−.21***
2. Age	.05	—	.01	−.08	.07	.00	.07	−.07	−.02	.10
3. Maternal education	.41***	.08	—	.26***	−.05	−.32***	.07	−.09*	.11**	−.21***
4. Race (White)	.29***	−.01	.25***	—	−.01	−.49***	−.03	−.01	.07	−.19***
5. Gender (male)	.02	.13**	−.08	−.03	—	.05	.01	−.03	.07	.23***
6. School demographic risk	−.33***	−.02	−.35***	−.46***	.05	—	.03	−.14***	−.07	.16***
7. Basic skills instruction	−.04	−.02	−.02	−.06	−.05	−.03	—	−.20***	.01	.07
8. Inference skills instruction	.03	.14***	.01	.02	.01	−.07	−.10*	—	−.02	−.05
9. Teacher–student closeness	.02	.00	.13***	.12**	−.14***	−.07	−.06	−.08	—	−.41***
10. Teacher–student conflict	−.26***	.03	−.28***	−.28***	.19***	.27***	.04	−.04	−.35***	—

Note. Third-grade correlations are contained in cells above the diagonal; fifth-grade correlations are contained in cells below the diagonal.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

Table 4

Repeated Measures Analyses of Math Test Scores (n = 587)

		B (SE) for Model 1				B (SE) for Model 2		
Predictor	***F* (*df*)**	**Low early skill**	**Average early skill**	**High early skill**	***F* (*df*)**	**Low early skill**	**Average early skill**	**High early skill**
Early skill group (*M/SE*)	70.0^{***} (2,374)	489.2_a (1.13)	501.4_b (0.47)	506.7_c (0.87)	7.44^{***} (2,342)	490.2_a (1.23)	501.4_b (0.46)	506.8_c (0.96)
Age	$1{,}153^{***}$ (1,374)				47.2^{***} (1,342)			
Group × Age	1.29 (2,374)	8.37 (0.57)	8.37 (0.24)	9.14 (0.43)	0.58 (2,342)	12.39 (1.96)	12.84 (1.88)	13.3 (1.89)
Demographic characteristics								
Race (White)[a]	5.75^{*} (1,374)	2.35 (0.98)	2.35 (0.98)	2.35 (0.98)	1.45 (1,342)	1.27 (1.05)	1.27 (1.05)	1.27 (1.05)
Gender (male)[a]	6.49^{*} (1,374)	1.92 (0.75)	1.92 (0.75)	1.92 (0.75)	8.85^{**} (1,342)	2.27 (0.76)	2.27 (0.76)	2.27 (0.76)
Maternal education[a]	52.49^{***} (1,374)	1.26 (0.17)	1.26 (0.17)	1.26 (0.17)	42.4^{***} (1,342)	1.15 (0.18)	1.15 (0.18)	1.15 (0.18)
Site	5.73^{***} (9,374)				5.73^{**} (9,342) *			
Classroom/teacher factors								
Basic skills instruction					2.84^{+}(1,342)			
Basic × Group					0.50 (2,342)	0.77 (0.53)	0.43 (0.30)	0.09 (0.45)
Basic × Age[a]					0.50 (1,342)	0.02 (0.03)	0.02 (0.03)	0.02 (0.03)
Inference skills instruction					1.35 (1,342)			
Inference × Group					2.85^{+}(2,342)	2.72_a (1.30)	-0.63_b (0.56)	−0.02 (1.06)
Inference × Age[a]					0.19 (1,342)	0.03 (0.06)	0.03 (0.06)	0.03 (0.06)
Teacher–student closeness					0.11 (1,342)			
Closeness × Group					1.65 (2,342)	0.20 (0.14)	0.01 (0.07)	−0.14 (0.14)
Closeness × Age[a]					4.73^{*} (1,342)	−0.11 (0.05)	−0.11 (0.05)	−0.11 (0.05)
Teacher–student conflict					3.57^{+} (1,342)			
Conflict × Group					0.28 (2,342)	−0.16 (0.13)	−0.18 (0.13)	−0.06 (0.15)
Conflict × Age[a]					2.86^{+} (1,342)	−0.08 (0.05)	−0.08 (0.05)	0.08 (0.05)
Basic × Closeness					4.67^{*} (1,342)			
Basic × Closeness × Group					0.67 (2,342)	−0.03 (0.01)	−0.01 (0.01)	−0.01 (0.01)
Inference × Closeness					0.66 (1,342)			

		B (SE) for Model 1				**B (SE) for Model 2**		
Predictor	***F (df)***	**Low early skill**	**Average early skill**	**High early skill**	***F (df)***	**Low early skill**	**Average early skill**	**High early skill**
Inference × Closeness × Group					1.01 (2,342)	−0.04 (0.03)	0.01 (0.01)	0.01 (0.03)
Basic × Conflict					0.02 (1,342)			
Basic × Conflict × Group					0.89 (2,342)	−0.00 (0.01)	−0.01 (0.01)	0.01 (0.02)
Inference × Conflict					4.49* (1,342)			
Inference × Conflict × Group					3.20* (2,342)	−0.10 (0.04)	0.01 (0.02)	−0.02 (0.03)
School demographic risk					3.24+			
School Risk × Age[a]					0.95	−3.16 (5.80)	−3.16 (5.80)	3.16 (5.80)
School Risk × Group					0.25 (2,342)	−5.95 (3.99)	−2.92 (2.20)	−2.79 (4.26)

Note. Age interactions were dropped from the model because they were nonsignificant. In the first row, means with different subscripts differed significantly ($p < .05$) according to F tests. The degrees of freedom were 372 for Model 1 and 340 for Model 2.

[a]Parameters were estimated across groups, so the same parameter is reported for each group.

+ $p < .10$.

* $p < .05$.

** $p < .01$.

*** $p < .001$.