

Promoting Reading Achievement and Countering the “Fourth-Grade Slump”: The Impact of Direct Instruction on Reading Achievement in Fifth Grade

Jean Stockard

National Institute for Direct Instruction and University of Oregon

Previous research has documented a substantial decline of standardized test scores of children from low-income backgrounds, relative to more advantaged peers, in later elementary grades, the so-called “fourth-grade slump.” This article examines changes in reading achievement from first to fifth grade for students in a large urban school system with a high proportion of students from economically deprived backgrounds. Students received first-grade reading instruction from Direct Instruction (DI), Open Court, or a mixture of reading curricula. Results indicate that students in schools using DI had significantly greater gains in both reading vocabulary and comprehension than students in the two other settings and that their average levels of achievement in fifth grade were above the national norms, thus countering the fourth-grade slump.

An extensive social science literature, spanning at least four decades, has documented the relationship between academic achievement, educational aspirations, eventual educational attainment, and adult occupational and economic success. Young people who do poorly in school are less likely to finish high school or attend college and have lower occupational status and lower incomes as adults (e.g., Blau & Duncan, 1967; Farkas, 1996; Jencks, Crouse, & Mueser, 1983; Sewell, Haller, & Ohlendorf, 1970). A slightly smaller, but still convincing, literature links early academic problems to later academic achievement, indicating that children with academic problems early in their school career are much more likely to also exhibit problems in later years (e.g., Juel, 1988). Reading is generally cited as the most important and central skill in promoting academic achievement, primarily because learning to read is basic to acquiring other skills and proceeding successfully through a schooling career (Murphy, 2004).

These associations carry special weight for children from low-income backgrounds, who often enter school with substantially lower levels of prereading skills or “readiness” to achieve. Moreover, even when low-income children approach the achievement levels of their more advantaged peers in the early grades, differences between them often widen in the later elementary period. This has come to be known as a “fourth-grade slump,” a substantial decline of standardized test scores of children from low-income backgrounds relative to their more economically advantaged peers as they progress past third grade and into the later elementary

years (Chall & Jacobs, 2003; Chall, Jacobs, & Baldwin, 1990; Hirsch, 2003; Rosenshine, 2002). In other words, although learning to read is important for later success for all children, those from economically deprived backgrounds appear to face special challenges in maintaining higher levels of academic achievement that can translate into adult occupational and economic success.

Reflecting the centrality of reading to long-term success, a large and growing body of literature has examined the impact of various curricula on reading achievement in the early elementary grades. Additional work has looked at their impact among different populations of older children and young adults. Surprisingly, however, there appears to be far less attention to the longer-term impact of various curricula on children's reading achievement, especially the changes in achievement across time periods that span the time of the fourth-grade slump. This article addresses that gap in the literature by examining changes in reading achievement from first to fifth grade for students in a large urban school system with a high proportion of students from economically deprived backgrounds. We compare the achievement of students who received first-grade reading instruction in three different settings: (a) schools that used Direct Instruction (DI), a curriculum that directly addresses areas seen as crucial in counteracting the fourth-grade slump; (b) schools using Open Court, another highly rated curriculum; and (c) schools using a variety of reading curricula. Our results indicate that students in schools with the DI curriculum had significantly greater gains in both reading vocabulary and comprehension than students in the two other settings and that their average levels of achievement in fifth grade were above the national norms, thus countering the fourth-grade slump.

BACKGROUND

A large body of literature has examined curricular and school reform models that can enhance student achievement. This literature has documented the importance of systematic and explicit instruction in promoting reading achievement (Adams, 1990; Anderson, Hiebert, Scott, & Wilkinson, 1985; Baker, Kameenui, Simmons, & Stahl, 1994; Bond & Dykstra, 1967; Chall, 1967; Foorman, 1995; Fukkink & deGlopper, 1998; Grossen, 1997; Institute of Child Health and Human Development, 2000; Juel & Minden-Cupp, 2000; Murphy, 2004; National Institute of Child Health and Human Development, 1996; National Reading Panel, 2000; Pflaum, Walberg, Karigianes, & Rasher, 1980; Smith et al., 2001; Snider, 1990; Snow, Burns, & Griffin, 1998; Stanovich, 1994). Meta-analyses examining specific curricula support this conclusion, showing that programs that embody these elements consistently result in larger achievement gains (Adams & Engelmann, 1996; American Federation of Teachers, 1998; Beck & McCaslin, 1978; Borman, Hewes, Overman, & Brown, 2003; Herman et al., 1999).

The Effectiveness of Direct Instruction

One of the most prominent explicit instructional approaches is Direct Instruction (DI; distinguished from other "direct instruction" approaches by the use of capital letters), which was developed by Siegfried Engelmann, Wesley Becker, and their colleagues (Engelmann, 2007; Engelmann & Carnine, 1982). DI curricula are specifically designed to accelerate students' learning by teaching more than traditional programs in the same amount of time. Unlike

many curricula, the DI programs are extensively field-tested before dissemination to ensure that they produce the greatest learning in the most efficient manner. The programs, which are commercially available through SRA/McGraw Hill and Sopris West, involve scripted lessons designed to provide teachers with the most effective wording to allow them to present tasks to students at a relatively high rate of speed. The amount of new material introduced in each lesson is carefully controlled, with applications becoming increasingly complex and designed so that, at the end of each lesson, all children will have mastered all of the content in the current lesson. The content of the lessons is also carefully designed to provide the basis for continued academic growth and understanding. Analyses of the DI curriculum suggest that, unlike traditional teaching methods, including those often termed “direct,” the DI approach teaches an underlying order of knowledge and provides the basis for accelerated cognitive growth (Carnine, Grossen, & Silbert, 1992).

Numerous studies have documented the superiority of DI in promoting reading achievement. These results have appeared with the general student population (e.g., Becker & Carnine, 1980; O’Brien & Ware, 2002; Stebbins, St. Pierre, Proper, Anderson, & Cerva, 1977; Vitale & Joseph, 2008) and with students with disabilities. Studies have demonstrated the effectiveness of DI reading programs with students with learning disabilities (Benner, 2007; Benner, Kinder, Beaudoin, Stein, & Hirschmann, 2005; Cooke, Gibbs, Campbell, & Shalvis, 2004; Kuder, 1990, 1991; Malmgren & Leone, 2000; Scarlato & Asahara, 2004), students with intellectual disabilities (Flores, Shippen, Alberto, & Crowe, 2004; Haring & Krug, 1975; Maggs & Morath, 1976; Malmgren & Leone, 2000; Riepl, Marchand-Martella, & Martella, 2008), children who demonstrate developmental delays (Flores & Ganz, 2007; Riepl et al., 2008), and students identified with emotional disturbance (Benner, 2007; Cook et al., 2004; Malmgren & Leone, 2000; Scarlato & Asahara, 2004; Strong, Wehby, Falk, & Lane, 2004). Additionally, DI reading programs have been effectively implemented outside of traditional elementary schools in a variety of settings, including middle schools (Dowdell, 1996; Grossen, 2004; Lewis, 1982; Shippen, Houchins, Steventon, & Sartor, 2005), high schools (Harris, Marchand-Martella, & Martella, 2000; Marchand-Martella, Martella, Orlob, & Ebey, 2000), a residential treatment center (Scarlato & Asahara, 2004), alternative schools (Steventon & Fredrick, 2003), and juvenile corrections facilities (Drakeford, 2002; Houchins, Jolivet, Krezmien, & Baltodano, 2008; Malmgren & Leone, 2000;). Last, DI programs have been shown to be effective in increasing reading achievement with English Language Learners (Grossen, 2004; Gunn, Smolkowski, Biglan, & Black, 2002; Kamps et al., 2007). As would be expected, studies have also found that the magnitude of the results is stronger when the program is implemented with greater fidelity (e.g., Ross et al., 2004; Stockard, 2009).

Longitudinal Studies

Numerous studies of DI, such as those previously noted, have documented its success in promoting achievement over a relatively short period of time, such as one to two years, and with children at various grade levels and ages. Yet far fewer have examined longer time periods. This lack is especially striking given the claim of Carnine et al. (1992) that DI should promote long-term cognitive growth, thus potentially countering the well documented decline in achievement for low-income children, relative to higher income peers in the later elementary grades.

After an extensive review of the literature, we found only one study, published over 25 years ago, that examined the impact of receiving DI in early elementary school on students' reading achievement in fifth and sixth grade, the later years of elementary school.¹ We also found a few other studies that examined somewhat shorter time spans.

Becker and Gersten's (1982) analysis of multiple implementations. Becker and Gersten (1982) examined the achievement of low-income fifth- and sixth-graders in five diverse sites, all of whom had completed grades 1 through 3 in DI as part of the large, federally funded Follow Through program. Students' achievement scores on the Wide Range Achievement Test (WRAT) and the Metropolitan Achievement Test were compared to those of students in local, demographically similar comparison groups who had received traditional educational programs in their early elementary years. Multivariate statistics were used to control for a number of variables that could affect achievement, including ethnicity, family income, home language, mother's education, gender, and number of siblings. Analyses of covariance that incorporated the control variables were conducted separately for each of the communities. Analyses involved data gathered in two different years, and across the subdimensions, as well as the total scores of the two achievement tests, thus providing multiple replications of the analyses. Three meta-analytic techniques were used to summarize the results, and all supported the conclusion that children who had received DI in the early grades had significantly higher achievement in fifth and sixth grade than students in the comparison groups. Becker and Gersten noted that the strongest and most consistent findings occurred with reading decoding, as measured by the WRAT, with a median effect size across years and sites of .48 (calculated from Table VI, Becker & Gersten, 1982, p. 83). More variable results occurred with the measure of word knowledge and total scores for reading.

The DI Follow Through program that Becker and Gersten (1982) examined ended when the children were in third grade. Even though the strong start that the DI children received resulted in reading achievement scores in the higher elementary grades that were significantly greater than those of demographically similar students in the comparison groups, their achievement relative to national norms declined after third grade. As Becker and Gersten (1982) described it:

2 years after the program had ended, all samples made appreciable, significant drops against the national norm group in . . . reading. . . . Though in many domains Follow Through graduates outperform the control students in Grades 5 and 6, . . . low-income Follow Through students are losing against the national normal sample. (p. 88)

This loss against the national norms parallels the fourth-grade slump noted in the literature as common among students from low income backgrounds (e.g., Chall et al., 1990; Hirsch, 2003).

Becker and Gersten (1982) concluded their article by noting a need to extend the principles of DI, including mastery learning, high levels of feedback, and incremental steps of instruction and learning, to the intermediate grades. In the decades following the publication of this work, DI curriculum developers, including Becker, heeded this call, focusing especially on ways to close the gap in vocabulary knowledge of low-income children and their higher income peers. The DI corpus has expanded to include instruction for students at higher grade levels and with a variety of higher-level subjects ranging from literary analysis to chemistry to legal reasoning.

¹We were also unable to find such longitudinal studies of other reading curricula.

In addition, the curriculum from grade 3 to 5 includes substantial content related to science and literature. Evaluations of these curricula indicate that they can successfully promote student learning at those levels (Carnine, 1991; Carnine et al., 1992; Carnine & Kameenui, 1992). We have not, however, been able to find any longitudinal studies that have examined changes in achievement through the elementary years when exposed to the contemporary DI curriculum.

Studies of single districts and shorter time periods. Several studies have examined the impact of receiving DI over shorter time periods than that addressed by Becker and Gersten (1982). For instance, Carlson and Francis (2002) compared the third-grade achievement of students receiving the DI program, *Reading Mastery*, in earlier grades with the achievement of students in demographically similar control schools who had used a traditional curriculum. Using multivariate analyses and a variety of achievement measures, they found that students with more exposure to DI had significantly higher achievement at the end of third grade, but that the largest increases relative to the comparison group were made during first grade. Similarly, Kamps and associates (2003) compared the impact of three different curricula—*Reading Mastery*, *Success for All*, and a literature based program—on growth in reading achievement over three years for students in the early primary grades. They found that students in the *Reading Mastery* curriculum had the greatest growth, followed by those having *Success for All*.

Two previous studies examined the sample used for the analysis in this article and focused on a subgroup of the students studied here. Mac Iver and associates (Mac Iver & Kemper, 2002; Mac Iver, Kemper, & Stringfield, 2003) used a subgroup of the schools studied in this article to examine changes in achievement for two cohorts of students exposed to DI or to a control curriculum. One cohort began DI in kindergarten and was followed through third grade, and the other cohort began DI in second grade and was followed through fifth grade. Results indicated positive impacts of DI on vocabulary and oral reading fluency, but less impact on reading comprehension. Addison and Yakimowski (2003) examined descriptive data regarding reading achievement from first through fifth grade for a portion of the schools in our analysis and data regarding achievement from first through fourth grade for another subset of the schools that we examined. They also used multivariate analyses to examine changes in achievement from kindergarten to second grade and from third to fifth grade for two subgroups of students. The descriptive results over the longest time period (first through fifth grade) indicated substantial differences in favor of DI. The differences were much smaller and often not significant for analyses of the shorter time periods.

Summary and Hypotheses

Even though a large body of literature has documented the relationship of DI to higher reading achievement for diverse groups of students, relatively little research has examined the relationship of DI to changes in achievement from first grade to the end of elementary school, a time period that is especially important in predicting later academic success. The major exception is work by Becker and Gersten (1982), published more than a quarter of a century ago. They found that students who had received DI through grade 3 had significantly higher reading achievement in grades 5 and 6 than demographically similar students in traditional curricula. They also noted, however, that the achievement of students in their study, most of whom were

from lower income families, was declining relative to national norms as they progressed to the higher grades, a finding that parallels the well documented fourth-grade slump.

Our analysis returns to the question posed by Becker and Gersten (1982), looking at the impact of DI on achievement gains from first to fifth grade. We approach this issue by looking at the relationship of exposure to the DI curriculum in early elementary school to reading achievement at the end of elementary school (grade 5) in a low-income urban school system. Specifically, we compare the changes in reading achievement from first to fifth grade of students exposed to three different curricular programs in first grade: (a) the highly rated DI program; (b) Open Court, another highly rated program; and (c) varied traditional instructional reading programs. If DI is more effective than other programs in promoting achievement, as Becker and Gersten found, we would expect that the changes from first to fifth grade would be larger for the students exposed to DI than for students in the two other groups. In addition, as noted previously, the DI curriculum has expanded markedly since the publication of Becker and Gersten's (1982) work to incorporate elements that should promote the development of higher-order cognitive skills and enhance continuing achievement. To the extent that these modifications and extensions have been successful, we expect that the students exposed to DI would have achievement scores that were closer to the national norms than students in the control groups.²

METHODOLOGY

The data for this analysis come from the Baltimore City Public School System (BCPSS) and regard the reading achievement of students who were fifth graders in either 2001–2002 or 2002–2003 and had been in the same schools five years earlier (in 1997–1998 or 1998–1999). The following sections describe the procedures, sample, measures, and analysis techniques that were used.

Procedures

BCPSS is similar to many other large city school districts that serve students with high levels of poverty and struggle with low levels of achievement. During the period from which our data were obtained, the average school within the system was 84% African American, 1% Hispanic, 14% non-Hispanic White, and had 75% of its students receiving free or reduced-price lunch.

In the late 1990s, curricular reforms were implemented in BCPSS elementary schools to address low achievement. Sixteen BCPSS schools chose to use DI for reading instruction. The curriculum was introduced as part of a whole school reform effort called the Baltimore Curriculum Project and sponsored by the Abell Foundation. The other schools followed BCPSS curricular guidelines to increase student achievement (Berkeley, 2002; Mac Iver, 2004; Mac Iver & Kemper, 2002; Stringfield & Yakimowski-Srebniak, 2005).

²Even though we examined data from the same school system as Mac Iver and Kemper (2002), and Addison and Yakimowski (2003), our work differs from theirs by using multivariate statistics to examine changes from first to fifth grade for all of the students exposed to DI in the system, rather than a subsample, and by comparing the impact of exposure to three different curricular programs in first grade.

Direct instruction. The major DI reading curriculum is *Reading Mastery Classic*, a scripted, mastery-based core reading program that focuses on decoding and comprehension. Students are placed in homogeneous groups according to skill level. Ideally, the teacher ensures that all members of the group achieve mastery on all material the program introduces. Students who master content substantially faster or slower than others in their group are placed into other groups in which students have skill profiles similar to those of the incoming student. Student skill is continuously monitored, and problems of mastery are addressed to ensure all students are at a level commensurate with their current level of skill. Instruction is designed to elicit frequent oral student responses, which increase engagement and create a high rate of active responding. Student responses during independent seat work are also closely monitored and immediately remediated (Marchand-Martella, Slocum, & Martella, 2004).

The students in the DI schools also received language instruction for 30 minutes per day. The curricula used for this instruction were *Language for Learning*, *Language for Thinking*, and *Reasoning and Writing*. These are general knowledge programs that focus on oral language development. The DI curriculum extended through fifth grade includes instruction in U.S. history, writing, reasoning, and spelling. The reading programs also include extensive general knowledge material related to the sciences and social studies. In general, the DI curriculum is designed to provide students with background knowledge and underlying schemata that can help them address and understand content throughout the academic curriculum. As noted, many of these extensions of the curriculum were a response to the findings of Becker and Gersten (1982), who expressed concern regarding falling achievement in the later elementary years and called for the development of additional instructional modules addressing this issue.

At the beginning of implementation, the Baltimore Curriculum Project contracted with the National Institute for Direct Instruction (NIFDI) to provide implementation support through pre-service and inservice training, coaching, and technical assistance to all the DI schools. The developer of DI is affiliated with NIFDI, and the organization prides itself on strict fidelity to the DI model as it was originally validated through extensive field testing.

When a school implements the NIFDI model, instructional programs are phased in over several years. In the first year, language and reading programs are introduced in kindergarten and first grade. In subsequent years, the curriculum is added to the higher grades and other parts of the program, including mathematics, are implemented. To ensure appropriate placement of students, the NIFDI model requires homogeneous instructional groupings, with regrouping as often as four times a year, monitoring student progress through direct observations, in-program tests, records of lessons completed, and at least weekly check-ins.

A NIFDI Implementation Manager (IM) trains teachers, assistants, and coaches. The IM is typically on site a total of approximately 35 person days per year, working in classrooms with the teachers and presenting in-service sessions that address problems teachers are experiencing. There also are weekly conference calls that address current problems each classroom is experiencing in meeting projected performance gains. All teachers receive preservice training and coaching until they teach each program to a minimum adequate level of fidelity. Teachers continue to receive in-service coaching to improve implementation fidelity. During the second year, teachers who perform well are identified as coaches and are deployed to work with other teachers in the school. NIFDI-led support is usually phased out after three years, although the

Baltimore Curriculum Project continued to operate the DI schools as Charter Schools and to provide on-going consultation and coaching for teachers.³

Eight of the schools began implementing DI in 1996–1997, three in 1997–1998 and the remaining five in 1998–1999. In the analysis reported in the following, students were designated as belonging to the intervention group (DI) only if they had received this instruction in first grade. If they received DI in grades 2–5 but not in grade 1, as occurred for students who were in first grade in 1997 in the five schools that began implementation in 1998–1999, they were omitted from the sample.

Control schools. Before 1998, the other schools in BCPSS were free to use any curriculum program they desired. There was no districtwide structured reading program, and schools reportedly used a variety of instructional programs. Yet schools in BCPSS had substantially lower achievement than other schools in the state and there was extensive pressure for reform (Stringfield & Yakimowski-Srebnick, 2005).

In response to these pressures, BCPSS embarked on a concentrated and extensive reform process. As part of this process, the system adopted new citywide reading curricula in the fall of 1998, using Open Court Reading in kindergarten through second grade. Open Court is a phonics-based, highly structured program that has been favorably reviewed by the Florida Center for Reading Research (2004) and the Oregon Reading First Center (2004) as a core instructional program for Reading First. Students in third grade and beyond received instruction from a Houghton Mifflin series. BCPSS provided extensive professional development support in 1998–1999 for teachers as they began to use this new curriculum. At the same time, the system introduced other reforms such as lowering student/teacher ratios, expanding kindergarten programs to full-day, and expanding before-school, after-school, and summer school programs (Berkeley, 2002; Mac Iver et al., 2003; Stringfield & Yakimowski-Srebnick, 2005). Note that these changes also affected students in the DI schools.

In this article, we differentiate students in the control schools into two groups: (a) those who had diverse first grade reading instruction (the 1997–1998 first-grade cohort), and (b) those who had first-grade reading instruction in Open Court (the 1998–1999 cohort). Note, however, that students with the diverse instruction in first grade had Open Court in second grade and that all of the students in both sets of control schools had a Houghton Mifflin series beginning in third grade.

Participants

Data for this study included student cohorts who began first grade in BCPSS in 1997–1998 or 1998–1999 and were in the same schools in fifth grade in either 2001–2002 (for the first graders in 1997–1998) or 2002–2003 (for the first graders in 1998–1999).⁴ Students who were held back

³Five of the 16 schools that introduced DI ceased their association with NIFDI soon after implementation and began to work with another provider of implementation support. To maintain an adequate sample size, the results for all 16 schools are combined for this analysis. Separate results for these two subsets of the DI schools are available upon request from the author and do not alter the substantive findings reported here.

⁴To preserve degrees of freedom and keep the sample size as large as possible, the two cohorts of DI students were grouped together for the analysis presented here. Results are virtually identical when they are separated. A methodological appendix to this article (Appendix A) provides descriptive results separating the two cohorts and details of the multivariate analyses separating the cohorts are available upon request from the author.

were not included in the analyses.⁵ There was substantial student turnover during the study period. Of the over 17,000 students in first grade in BCPSS in 1997–1998 and 1998–1999, about half (49.7%) had left BCPSS five years later, whereas almost one-quarter (23.3%) had transferred to another school within the system. Slightly more than one-fourth of the students who were in first grade in the system in 1997–1998 and 1998–1999 were in fifth grade in the same school five years later. Students in the DI schools were slightly, but significantly, more likely to remain in the same schools in fifth grade than those in the two control groups. Almost 30% (29.3%) of the students in the DI schools persisted in the same schools, compared to 26% of those in the other two groups ($\chi^2 = 7.63$, $df = 2$, $p = .02$).

The panel sample (those that were in the same schools in fifth grade that they attended in first grade, five years earlier) analyzed in this report includes a total of 4,572 students. Of these, 500 were in DI Schools, 2,197 were in control schools with variable curricula in first grade, and the remaining 1,875 were in first grade in schools that implemented Open Court.⁶ Six schools in BCPSS, all within the control groups, had no students who persisted from first to fifth grade. As a result, the sample included students from 113 schools. A methodological appendix to this article (Appendix A) provides additional details on the number of students in these various groups in each year.

Instrumentation

The Comprehensive Test of Basic Skills/TerraNova (CTBS; CTB/McGraw-Hill, 2001), a widely used standardized achievement test, was administered to all students in the spring of each school year from 1997–1998 through 2002–2003 as part of a systemwide testing program. The fourth edition was administered in the spring of 1998 and 1999, and the fifth edition in 2002 and 2003. Two subtest scores, Reading Comprehension and Reading Vocabulary, were analyzed. Normal Curve Equivalent (NCE) scores were used to help ensure comparability from one year to another and to allow the use of statistical calculations. Because the meaning of NCE scores is not intuitively obvious, the scores were also converted, when appropriate for descriptive purposes, into percentiles using a standard conversion table.

To adjust for the impact of socioeconomic status on student achievement and achievement gains over time, we used, as a control variable, the proportion of students within a school who received free or reduced-price lunch (FRL). Because we did not have access to individual-level demographic data, a school-level measure was used as a proxy. In preliminary analyses, we also used a factor score incorporating measures of race/ethnicity and school

⁵Mac Iver and Kemper (2002) analyzed data from some of the schools included in this analysis and found that the DI schools had a lower rate of retention (“failing” a grade) than other schools in the system. This could conceivably result in a larger proportion of potentially “low achieving” or “higher risk” students in the fifth grade in the DI schools and thus provides a conservative test of the impact of DI on changes in achievement from first to fifth grade. However, when students who were in first grade in 1997 and were in fifth grade in 2003 (six years after first grade, rather than five) were included in the multivariate analyses, substantive results were virtually identical to those reported in this article. These results are also available upon request from the author.

⁶There was no way to control for the possibility that a student had attended multiple other schools between first and fifth grade. It is possible that some of the students in the panel sample were in the targeted schools in both first and fifth grade, but had attended other schools in the interim. Assuming that such children would have less of the “full treatment,” this would bias results in a conservative direction.

poverty as a control variable. Results, available from the author, were identical to those reported here.

Most schools in BCPSS included high proportions of students receiving FRL. The average student in the sample attended a school where 71% of students were in this category. Students in the panel group differed from other first-grade students (those who did not remain in the same schools from first to fifth grade) in the average socio-demographic characteristics of their schools, with those who persisted being significantly more likely to attend higher status schools (72% FRL vs. 76% FRL, $F = 226.275$, $p < .001$). At the same time, students in the DI schools were more likely than other students to be in high-poverty schools (75% for students in DI schools vs. 71% for other students; $F = 68.983$, $p < .001$). There was, however, no significant interaction; that is, the differences in school poverty context between students in the panel group and the other students was the same across the three groups of schools (DI schools and the two control schools; F interaction of group and panel status = 0.81, $p = .44$).

In addition, as could be expected, the students who persisted in the same schools had higher first-grade achievement than their more mobile counterparts. These differences were similar across the three groups for the measure of comprehension (F interaction = 1.89, $p = .15$), but varied across the groups for the measure of vocabulary. With this measure, the difference between the persisters (those in the panel group) and the mobile students was smaller for students in the DI schools than for students in the other schools (F interaction, = 7.72, $p < .001$). (See Appendix A for full descriptive data.)

Note that these differences resulted in the students in the DI schools being somewhat more at risk in first grade than the other students. They attended schools in which more students received FRL and they also had lower average vocabulary scores in first grade. If anything, this could produce a conservative impact on the results (a smaller chance of significant findings in favor of DI).

Even though students who remained in the panel sample and were included in our analysis were more likely than those who left the school system to be in schools with more advantaged characteristics, it is important to stress that the average student in the sample attended a very high-poverty school with substantial proportions of racial/ethnic minorities in attendance. In fact, the average (mean) values previously given understate the degree of segregation and poverty in BCPSS. A somewhat more accurate picture may come from positional measures, such as quartiles. These indicate that 50% of the students attended schools where at least 97% of the students were African American, and 75% attended schools where at least 78% of the students were African American. Similarly, 75% of the students attended schools where two-thirds of the students were on FRL.

Analysis

The impact of DI on changes in students' reading achievement from first to fifth grade was examined in two ways. First, we used descriptive statistics, paired t -tests, and repeated measures analysis of variance to examine the average change in students' scores over the years in each of the three groups of schools (DI, varied first grade curricula, and Open Court in first grade). The percentage change in NCE scores and effect sizes were also used to assess the substantive magnitude of these changes. Effect sizes were calculated using the method developed by Dunlap,

Cortina, Vaslow, and Burke (1996) for calculating Cohen's d with correlated samples.⁷ If having DI as a core curriculum enhances growth in students' achievement from first to fifth grade more than either the varied curricula or Open Court, we would expect a significant interaction effect in the analysis of variance and that the pairwise comparisons would indicate that students in DI schools had stronger growth in achievement over time.

Second, mixed-model regressions were conducted that regressed fifth grade achievement on first-grade achievement, the percentage of students in the school who received FRL, and dummy variables for group. Students who received Open Court were the omitted category in the dummy variable coding. To the extent that having DI in first grade was more effective at promoting higher student achievement in later years, it would be expected that students in the DI schools would have higher fifth-grade achievement scores than those in the control schools, even when school characteristics and first-grade achievement were controlled.

Mixed models are particularly appropriate for analyzing multilevel data, such as that regarding students and the schools that they attend. In these models a "random variable" is used to control for differences between schools (often termed the Level 2 entity) while calculating regression coefficients regarding the impact of variables from both students and schools on achievement. The random variable is equivalent to having a separate intercept in the regression equation for each school. The coefficients associated with the various individual and school-related variables are then calculated while this between school variance is controlled. The analysis also allows one to calculate the amount of variance in the dependent variable that occurs between schools (Raudenbush & Bryk, 2002; Singer, 1998). We used the mixed procedure in SAS for these calculations.

RESULTS

Descriptive Statistics and Analysis of Variance

Descriptive results regarding the first and fifth-grade reading achievement of students in the three groups of schools are summarized in Tables 1 through 4. The mean and standard deviation of NCE scores for each measure and grade and the associated percentage change in these scores are in Table 1, the results of paired t -tests and associated effect sizes are in Table 2, the percentiles that correspond to the means in Table 1 are in Table 3, and the results of the repeated measures analyses of variance are in Table 4. The percentiles can be interpreted as the score that an average student in each group would have in a given grade. All calculations were done with NCE scores.

As hypothesized, the students in the DI schools had significantly greater gains in achievement from first to fifth grade than students in the other two groups. This is indicated by the significant interaction effects in the analyses of variance (Table 4), the larger average differences and effect sizes for the DI group (Table 2), and the larger percentage changes for the DI group (Table 1). In first grade, students in the DI schools had lower average vocabulary and comprehension scores than students in either of the other two groups of schools. The average student in the DI schools

⁷McLean, O'Neal, and Barnette (2000) cautioned that effect sizes calculated with NCE scores are inherently smaller than those with other metrics. Thus, the effect sizes presented might be a conservative estimate.

TABLE 1
Descriptive Statistics, First and Fifth Grade Norm Equivalent Reading Achievement Scores,
NIFDI-Supported Schools, Other DI Schools, and Control Schools, BCPSS, 1998–2003

	<i>Direct Instruction</i>		<i>Variable Curricula</i>		<i>Open Court</i>	
	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Vocabulary						
First	43.6	22.0	47.3	21.9	54.4	20.7
Fifth	51.8	19.7	49.0	17.8	50.5	16.8
Percent change	18.8		3.7		-7.0	
Comprehension						
First	42.5	19.8	43.4	20.1	47.4	20.0
Fifth	51.6	18.1	49.5	17.1	50.0	16.3
Percent change	21.6		14.1		5.6	

Note. NIFDI = National Institute for Direct Instruction; DI = Direct Instruction; BCPSS = Baltimore City Public School System.

TABLE 2
Average Differences, Correlations, Paired (Dependent Sample) *t*-Scores, and Effect Sizes

	<i>Average Difference</i>	<i>r</i>	<i>t</i>	<i>p</i>	<i>N</i>	<i>Effect Size</i>
Direct Instruction						
Vocabulary	8.2	0.55	9.09	<.001	490	0.39
Comprehension	9.2	0.53	11.05	<.001	494	0.48
Variable Curricula						
Vocabulary	1.7	0.47	3.84	<.001	2,143	0.09
Comprehension	6.1	0.53	15.74	<.001	2,175	0.33
Open Court						
Vocabulary	-3.8	0.47	-8.38	<.001	1,841	-0.20
Comprehension	2.6	0.49	6.12	<.001	1,843	0.14

TABLE 3
Percentile Score of Average Student by School Type, First and Fifth Grade

	<i>Direct Instruction</i>	<i>Variable Curricula</i>	<i>Open Court</i>
Vocabulary			
First	38	45	58
Fifth	53	48	51
Comprehension			
First	36	38	45
Fifth	53	49	50

Note. The percentiles in Table 3 correspond to the average (*M*) Normal Curve Equivalent values given in Table 1. They may be interpreted as the percentile at which an average student in a given group and grade would score. Median values for all scores are also given in the Appendix to this article.

TABLE 4
Results of Repeated Measures Analysis of Variance

<i>Measurement</i>	<i>F</i>	<i>df</i>	<i>p</i>
Vocabulary			
Achievement (repeated)	29.73	1,4471	<.001
Group	37.33	2,4471	<.001
Group by achievement	81.06	2,4471	<.001
Comprehension			
Achievement (repeated)	315.77	1,4509	<.001
Group	9.76	2,4509	<.001
Group by achievement	32.04	2,4509	<.001

Note. Simple one-way analysis of variance comparing scores between the three groups for each measure and grade also indicate significant differences: for first grade vocabulary, $F = 75.991$, $p < .001$; for first grade comprehension, $F = 24.528$, $p < .001$; for fifth grade vocabulary, $F = 6.69$, $p = .001$; for fifth grade comprehension, $F = 2.98$, $p = .05$.

scored at the 38th percentile in vocabulary and the 36th percentile in comprehension, whereas the average student in the control schools scored at the 45th (variable curricula schools) or 58th percentile (Open Court schools) in vocabulary and the 38th (varied curricula) and 45th percentile (Open Court) in comprehension. By fifth grade, these same DI students had higher scores than students in the other schools and scored above the national mean on both measures.

The differences in average change from first to fifth grade were not small in magnitude. On average, students in the DI schools had vocabulary scores in fifth grade that were 19% higher than their scores in first grade, compared to an increase of 4% for those in the varied curricula first grades and a decline of 7% for the students receiving Open Court. Students in all the groups had higher comprehension scores in fifth grade than in first grade, but again the changes were substantially larger for those in the DI schools—an average increase of 22%—compared to changes of 14% for those with the varied first grade curriculum and 6% for those with Open Court. These conclusions are supported by the effect sizes, which surpass the .25 criterion for substantively significant effects (Wolf, 1986) for both measures for the students in the DI schools, but for only one of the other four comparisons (for comprehension for students with the varied first-grade curriculum).

Mixed-Model Results

Table 5 gives the results of the mixed-model analyses, regressing fifth-grade achievement on first-grade achievement and controlling for differences between the schools in poverty context. The first panel gives results regarding vocabulary, and the second gives results regarding comprehension. Four models were tested. Model 1 is the baseline “intercept only” or “random effects” model and only includes schools as a random variable. This model tests the null hypothesis that the schools are equal in average achievement. The associated statistics are reported in the lower lines of each panel of the table. The correlation ratio is the proportion of variance in the dependent variable that is between schools as opposed to between students. It can be seen that from 7% to 9% of the variance in achievement is between schools rather than simply between

TABLE 5
 Mixed Model Regressions, Fifth Grade Reading (Norm Equivalent Scores) on School Advantage/Disadvantage, First Grade Achievement, and School Type
 (DI, Varied First Grade Curriculum, and Open Court)

	Model 1				Model 2				Model 3				Model 4				
	b	SE	p	Change	b	SE	p	Change	b	SE	p	Change	b	SE	p	Change	df
Vocabulary																	
Intercept	49.45	0.59	<.0001		63.21	2.50	<.0001		64.17	2.49	<.0001		37.47	2.65	<.0001		
School % FRL	—	—	—		-18.70	3.33	<.0001		-19.20	3.32	<.0001		-12.10	3.41	0.0004		
DI	—	—	—		—	—	—		0.98	1.54	0.52		5.16	1.56	0.001		
Varied first grade	—	—	—		—	—	—		-1.55	0.54	0.004		1.30	0.48	0.007		
First grade achievement	—	—	—		—	—	—		—	—	—		0.40	0.01	<.0001		
Random effect (schools)	28.32	4.96	<.0001		19.89	3.80	<.0001		19.35	3.73	<.0001		22.11	4.05	<.0001		
Residual	283.85	6.03	<.0001		283.89	6.0352	<.0001		283.41	6.03	<.0001		215.97	4.63	<.0001		
-2log likelihood	38602.1				38574.8				38565.2				36913.6				
BIC	38616.1				38593.5				38593.2				36946.2				
Change in LL																	
Correlation ratio	0.09																1
Comprehension																	
Intercept	49.37	0.51	<.0001		62.10	2.11	<.0001		62.46	2.10	<.0001		36.65	2.23	<.0001		
School % FRL	—	—	—		-17.30	2.82	<.0001		-17.74	2.80	<.0001		-10.89	2.86	0.0001		
DI	—	—	—		—	—	—		1.37	1.32	0.30		3.27	1.32	0.01		
Varied first grade	—	—	—		—	—	—		-0.45	0.52	0.38		1.36	0.45	0.002		
First grade achievement	—	—	—		—	—	—		—	—	—		0.44	0.01	<.0001		
Random effect (schools)	20.16	3.68	<.0001		12.90	2.70	<.0001		12.41	2.65	<.0001		14.69	2.83	<.0001		
Residual	262.91	5.56	<.0001		262.96	5.57	<.0001		262.98	5.57	<.0001		191.19	4.08	<.0001		
-2log likelihood	38546				38514				38512				36654				
BIC	38560				38533				38540				36687				
Change in LL																	
Correlation ratio	0.07																1

Note. No coefficient is included for Open Court because it is the omitted category for the dummy variables. FRL = free and reduced-price lunch; DI = direct instruction; BIC = Bayesian information criterion. LL = -2log likelihood statistic.

students. The estimates, z values, and probabilities associated with the random effects test the null hypothesis that the variation between schools equals zero once variables in a model are controlled. These values associated with the residual test the null hypothesis that variation between individuals equals zero once the variables in the model and school differences are controlled. The null hypotheses can be easily rejected with all the models. There is significant variation between schools and also between students in all the models. This is as we would expect, for there are undoubtedly many factors that can influence student achievement in addition to those available for this analysis.

Model 2 adds the percentage of students receiving FRL within the school as a predictor, and results with this model indicate the extent to which the poverty context of students' schools is related to their achievement in fifth grade. As would be expected, the coefficients, with both dependent measures, indicate that students had significantly higher achievement in fifth grade when they attended schools with fewer low income students. Model 3 adds the dummy variables indicating the type of school that students attended, testing the hypothesis that students' achievement in fifth grade varied between the three groups of schools once school poverty context was controlled. Model 4 adds first-grade achievement, controlling for students' skills at the end of first grade.

The -2 log likelihood measures and the Bayesian information criterion (BIC) values can be used to examine the relative fit of the data to the models. Lower values indicate a better fit. Differences between the log likelihood measures have a chi-square distribution, and the comparisons between these values are in the final rows of each part of Table 2. The BIC values provide a descriptive summary of the fit of the models, with lower values indicating a better fit. Both the BIC values and the changes in the -2 log likelihoods indicate that Model 4, which includes all the variables, provides the best fit to the data.⁸

The coefficients associated with the dummy variables for group support the conclusions obtained with the descriptive results in Tables 1 through 4. They indicate that students who attended DI schools in first grade had significantly higher reading achievement scores in fifth grade than students in other types of schools, even when the poverty level of their school and their levels of first-grade achievement were controlled. The students in the varied curriculum first grades also had significantly higher scores than the students in the Open Court schools, but the differences were not as large. The mixed-model results indicate that, on average, once characteristics of their schools and their first-grade achievement were controlled, students in the DI schools would have vocabulary achievement scores in fifth grade that were more than five NCE points higher than those in the Open Court schools and comprehension achievement scores that were more than three points higher. The students in the schools with the varied curriculum would have scores from 1.3 to 1.4 NCE points higher than those in the Open Court schools.

SUMMARY AND DISCUSSION

Our results from a low-income large urban school district indicate that it is possible to promote higher reading achievement over time. The data presented indicate that the efforts of BCPSS

⁸We also tested interaction effects, specifically the interaction of group and school characteristics and the interaction of group and first-grade achievement. Neither of these sets of interaction effects was significant.

to promote extensive systemwide change and higher student achievement have met with success. With one exception (changes in vocabulary scores for students who had Open Court in first grade), students in all three groups of schools had significantly higher reading achievement in fifth grade than in first grade.

However, as hypothesized, our results also indicate that students who had the DI curriculum through their elementary grades had significantly greater gains than students in the other curricula. These differences were both statistically and substantively significant and appeared with both descriptive and inferential analyses, with strong controls for the socio-demographic context of the schools they attended, and controlling for levels of first-grade achievement.

In addition, the results suggest that the DI curriculum may help counteract the so-called fourth-grade slump, when students from low-income backgrounds begin to fall progressively farther behind their more advantaged peers. Even though the students in the DI schools had first-grade achievement scores that were lower, on average, than those of students in the other schools, by fifth grade they had significantly higher average achievement scores than students in the other schools; and these scores were above the national average. In contrast, only one of the four fifth-grade achievement measures for the control groups averaged above the 50th percentile (the national average), and this represented a decline from the first-grade score. Thus, our results suggest that the modifications and expansions of the DI curriculum that occurred after Becker and Gersten's (1982) work have been successful in promoting higher order cognitive skills and continuing achievement, at least through the end of fifth grade.

Hirsch (2003) summarized three basic principles that influence students' reading comprehension and are related to the decline in the later grades: fluency, allowing the student to focus on comprehension; vocabulary, which increases the probability of comprehending more complex material; and domain knowledge, a threshold level of understanding that makes it easier to understand ever more complex and demanding material. The students in the DI schools had significantly greater achievement gains on both the measure of vocabulary and the measure of comprehension than students in the other schools, reflecting at least two of the three areas that Hirsch highlighted. In addition, the achievement gains over time for the DI students were stronger for the measure of comprehension than vocabulary. Given the centrality of comprehension to learning in the upper grades, this difference can also be seen as supporting the efficacy of DI in counteracting declines in achievement in the upper elementary grades.

It is, of course, very important to replicate the work reported here, especially given the apparent paucity of longitudinal research. It would be important to examine changes in achievement over the elementary years and beyond with both the DI curriculum and other curricula. Our results suggest that students in the DI curriculum had higher levels of achievement than those with a variety of approaches or Open Court in first grade, a finding reported elsewhere (e.g., Crowe, Connor, & Petscher, 2009; O'Brien & Ware, 2002). It is possible, however, that once school reforms became more fully implemented and stabilized, the achievement of students in these other curricula would be higher; this possibility should be investigated.

It would also be important to have studies that include measures of achievement prior to beginning school. Our study examined changes from the primary to intermediate grade-school years (from the end of first grade to the end of fifth grade), as have other examinations of the fourth-grade slump, but we had no measure of entry-level skills prior to beginning school. Having such data could provide better estimates of the long-term impact of curricula on growth. In addition, a very worthwhile area of further research could be to investigate the way in which

extending academic DI into the preschool years might increase first-grade performance and result in even higher achievement at fifth grade for especially vulnerable children.

Finally, future analyses would be improved by including individual measures of socioeconomic characteristics. Given limitations of our data set, we did not have individual level measures of socio-demographic characteristics. Although substantial literature has documented the influence of school contexts of disadvantage on achievement, the most precise estimates of the effects found here would occur if we had both individual- and school-level measures of socio-demographic characteristics.

Earlier analyses of the impact of DI on student achievement in BCPSS, using subsets of the data examined here, recommended that the system continue using the program (Addison & Yakimowski, 2003; Mac Iver et al., 2003). Our results would confirm that recommendation and suggest, as have many other studies previously cited, that students are well served by the DI curriculum. Even more important, our results suggest that the curriculum has long-term impacts and, at least for students in this high-poverty school system, can help counter the well documented tendency for declining achievement over time. The importance of this result for the well-being of low income students, as well as for society as a whole, cannot be overstated (see Belfield & Levin, 2007).

ACKNOWLEDGMENTS

I thank Cristy Coughlin, Kurt Engelmann, Zigfried Engelmann, Daniel Johnston, Jerry Silbert, and Gary Davis for very helpful comments on earlier drafts of this article. Any errors remaining are my sole responsibility.

REFERENCES

- Adams, M. J. (1990). *Beginning to read: Thinking and learning about print*. Cambridge, MA: MIT Press.
- Adams, G. L., & Engelmann, S. (1996). *Research on Direct Instruction: 25 years beyond DISTAR*. Seattle, WA: Educational Achievement Systems.
- Addison, K. L., & Yakimowski, M. E. (2003). *An evaluation of the Direct Instruction program*. Baltimore, MD: Baltimore City Public School System, Division of Research, Evaluation, Assessment, and Accountability.
- American Federation of Teachers. (1998). *Building on the best, learning from what works: Seven promising reading and language arts programs*. Washington, DC: Author.
- Anderson, R. C., Hiebert, E. H., Scott, J. A., & Wilkinson, I. A. G. (1985). *Becoming a nation of readers: The report of the Commission on Reading*. Washington, DC: National Institute of Education.
- Baker, S. K., Kameenui, E. J., Simmons, D. C., & Stahl, S. A. (1994). Beginning reading: Educational tools for diverse learners. *School Psychology Review*, 23, 372–391.
- Beck, I. L., & McCaslin, E. S. (1978). *An analysis of dimensions that affect the development of code-breaking ability in eight beginning reading programs* (LRDC Report No. 1978/6). Pittsburgh, PA: University of Pittsburgh.
- Becker, W. C., & Carnine, D. W. (1980). Direct Instruction: An effective approach to educational intervention with the disadvantaged and low performers. In B. B. Lahey & E. E. Kazdin (Eds.), *Advances in clinical child psychology* (Vol. 3, pp. 429–473). New York, NY: Plenum.
- Becker, W. C., & Gersten, R. (1982). A follow-up of Follow Through: The later effects of the Direct Instruction model on children in fifth and sixth grades. *American Educational Research Journal*, 19, 75–92.
- Belfield, C. R., & Levin, H. M. (2007). *The price we pay: Economic and social consequences of inadequate education*. Washington, DC: Brookings.

- Benner, G. J. (2007). The relative impact of remedial reading instruction on the basic reading skills of students with emotional disturbance and learning disabilities. *Journal of Direct Instruction, 7*, 1–15.
- Benner, G. J., Kinder, D., Beaudoin, K. M., Stein, M., & Hirschmann, K. (2005). The effects of the “Corrective Reading Decoding” program on the basic reading skills and social adjustment of students with high-incidence disabilities. *Journal of Direct Instruction, 5*, 67–80.
- Berkeley, M. (2002). The importance and difficulty of disciplined adherence to the educational reform model. *Journal of Education for Students Placed at Risk, 7*, 221–239.
- Blau, P. M., & Duncan, O. D. (1967). *The American Occupational Structure*. New York, NY: Wiley.
- Bond, G., & Dykstra, R. (1967). The cooperative research program in first-grade reading instruction. *Reading Research Quarterly, 2*, 5–142.
- Borman, G. D., Hewes, G. M., Overman, L. T., & Brown, S. (2003). Comprehensive school reform and achievement: A meta-analysis. *Review of Educational Research, 73*, 125–230.
- Carlson, C. D., & Francis, D. J. (2002). Increasing the reading achievement of at-risk children through Direct Instruction: Evaluation of the Rodeo Institute for Teacher Excellence (RITE). *Journal of Education for Students Placed At Risk, 7*, 141–166.
- Carnine, D. (1991). Curricular interventions for teaching higher-order thinking for all students. Introduction to the special series. *Journal of Learning Disabilities, 24*, 261–269.
- Carnine, D., Grossen, B., & Silbert, J. (1992). Direct instruction to accelerate cognitive growth. *ADI News*, Fall, 33.
- Carnine, D., & Kameenui, E. (1992). *Teaching higher order thinking to all students*. Austin TX: Pro Ed.
- Chall, J. S. (1967). *Learning to read: The great debate*. New York, NY: McGraw Hill.
- Chall, J. S., & Jacobs, V. A. (2003). Poor children’s fourth-grade slump. *American Educator, 27*(1), 14–17.
- Chall, J. S., Jacobs, V. A., & Baldwin, L. E. (1990). *The reading crisis: Why poor children fall behind*. Cambridge, MA: Harvard University Press.
- Cooke, N. L., Gibbs, S. L., Campbell, M. L., & Shalvis, S. L. (2004). A comparison of Reading Mastery Fast Cycle and Horizons Fast Track A-B on the reading achievement of students with mild disabilities. *Journal of Direct Instruction, 4*, 139–151.
- Crowe, E. C., Connor, C. M., & Petscher, Y. (2009). Examining the core: Relations among reading curricula, poverty, and first through third grade reading achievement. *Journal of School Psychology, 47*, 187–214.
- CTB/McGraw-Hill. (2001). *TerraNova technical manual report*. Monterey, CA: Author.
- Dowdell, T. (1996). *The effectiveness of Direct Instruction on the reading achievement of sixth graders*. (ERIC Document Reproduction Service No. 396268).
- Drakeford, W. (2002). The impact of an intensive program to increase the literacy skills of incarcerated youth. *Journal of Correctional Education, 53*, 139–144.
- Dunlap, W. P., Cortina, J. M., Vaslow, J. B., & Burke, M. J. (1996). Meta-analysis of experiments with matched groups or repeated measures designs. *Psychological Methods, 1*, 170–177.
- Engelmann, S. (2007). *Teaching needy kids in our backward system: 42 years of trying*. Eugene, OR: ADI Press.
- Engelmann, S., & Carnine, D. (1982). *Theory of instruction: Principles and applications*. New York, NY: Irvington.
- Farkas, G. (1996). *Human capital or cultural capital: Ethnicity and poverty groups in an urban school district*. New York, NY: Aldine.
- Flores, M. M., & Ganz, J. B. (2007). Effectiveness of Direct Instruction for teaching statement inference, use of facts, and analogies to students with developmental disabilities and reading delays. *Focus on Autism and Other Developmental Disabilities, 22*, 244–251.
- Flores, M. M., Shippen, M. E., Alberto, P., & Crowe, L. (2004). Teaching letter-sound correspondence to students with moderate intellectual disabilities. *Journal of Direct Instruction, 4*, 173–188.
- Florida Center for Reading Research. (2004). *Open Court Reading Pre-K*. Tallahassee, FL: Author. Retrieved from <http://www.fcrr.org/index.htm>
- Foorman, B. R. (1995). Research on “the great debate”: Code-oriented versus whole language approaches to reading instruction. *School Psychology Review, 24*, 376–392.
- Fukkink, R. G., & deGlopper, K. (1998). Effects of instruction in deriving word meaning from context: A meta-analysis. *Review of Educational Research, 68*, 450–469.
- Grossen, B. (1997). *A synthesis of research on reading from the National Institute of Child Health and Human Development*. Eugene, OR: University of Oregon.

- Grossen, B. (2004). Success of a Direct Instruction model at a secondary level school with high-risk students. *Reading & Writing Quarterly*, 20, 161–178.
- Gunn, B., Smolkowski, K., Biglan, A., & Black, C. (2002). Supplemental instruction in decoding skills for Hispanic and non-Hispanic students in early elementary school. *Journal of Special Education*, 36, 69–79.
- Haring, N. G., & Krug, D. A. (1975). Evaluation of a program of systematic instructional procedures for extremely poor retarded children. *American Journal on Mental Retardation*, 79, 627–631.
- Harris, R. E., Marchand-Martella, N. E., & Martella, R. C. (2000). Effects of a peer-delivered Corrective Reading program. *Journal of Behavioral Education*, 10, 21–36.
- Herman, R., Aladjam, D., McMahon, P., Masem, E., Mulligan, I., Smith, O., et al. (1999). *An educator's guide to school-wide reform*. Washington, DC: American Institutes for Research.
- Hirsch, E. D., Jr. (2003). Reading comprehension requires knowledge—Of words and the world: Scientific insights into the fourth-grade slump and the nation's stagnant comprehension scores. *American Educator*, 27, 10–29.
- Houchins, D. E., Jolivet, K., Krezmien, M. P., & Baltodano, H. M. (2008). A multi-state study examining the impact of explicit reading instruction with incarcerated students. *Journal of Correctional Education*, 5, 65–85.
- Institute of Child Health and Human Development. (2000). *Report of the National Reading Panel. Teaching Children to Read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction* (NIH Publication No. 00-4769). Washington, DC: US Government Printing Office.
- Jencks, C., Crouse, J., & Mueser, P. (1983). The Wisconsin model of status attainment: A national replication with improved measures of ability and aspiration. *Sociology of Education*, 56, 3–19.
- Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology*, 80, 437–447.
- Juel, C., & Minden-Cupp, C. (2000). Learning to read words: Linguistic units and instructional strategies. *Reading Research Quarterly*, 35, 458–492.
- Kamps, D. M., Abbott, M., Greenwood, C., Arreaga-Mayer, C., Wills, H., Longstaff, J., et al. (2007). Use of evidence-based, small group reading instruction for English language learners in elementary grades: Secondary-tier intervention. *Learning Disability Quarterly*, 30, 153–168.
- Kamps, D. M., Will, H. P., Greenwood, C. R., Thorne, S., Lazo, J. F., Crockett, J. L., et al. (2003). Curriculum influences on growth in early reading fluency for students with academic and behavioral risks: A descriptive study. *Journal of Emotional and Behavioral Disorders*, 11, 211–224.
- Kuder, S. J. (1990). Effectiveness of the DISTAR reading program for children with learning disabilities. *Journal of Learning Disabilities*, 23, 69–71.
- Kuder, S. J. (1991). Language abilities and progress in a Direct Instruction reading program for students with learning disabilities. *Journal of Learning Disabilities*, 24, 124–127.
- Lewis, A. (1982). An experimental evaluation of a direct instruction programme (Corrective Reading) with remedial readers in a comprehensive school. *Educational Psychology*, 2, 121–135.
- Mac Iver, M. A. (2004). Systemic supports for comprehensive school reform: The institutionalization of Direct Instruction in an urban school system. *Journal of Education for Students Placed At Risk*, 9, 303–321.
- Mac Iver, M. A., & Kemper, E. (2002). The impact of direct instruction on elementary students' reading achievement in an urban school district. *Journal of Education for Students Placed At Risk*, 7, 197–220.
- Mac Iver, M. A., Kemper, E., & Stringfield, S. (2003). *The Baltimore curriculum project: Final report of the four-year evaluation study*. Baltimore, MD: Center for Research on the Education of Students Placed at Risk, Johns Hopkins University.
- Maggs, A., & Morath, P. (1976). Effects of direct verbal instruction on intellectual development of institutionalized moderately retarded children: A 2-year study. *Journal of Special Education*, 10, 357–364.
- Malmgren, K. W., & Leone, P. E. (2000). Effects of a short-term auxiliary reading program on the reading skills of incarcerated youth. *Education and Treatment of Children*, 23, 239–247.
- Marchand-Martella, N., Martella, R., Orlob, M., & Ebey, T. (2000). Conducting action research in a rural high school setting using peers as Corrective Reading. *Rural Special Education Quarterly*, 19, 20–30.
- Marchand-Martella, N. E., Slocum, T. A., & Martella, R. C. (2004). *Introduction to Direct Instruction*. Boston, MA: Pearson.
- McLean, J. E., O'Neal, M. R., & Barnette, J. J. (2000, November). *Are all effect sizes created equal?* Paper presented at the annual meeting of the Mid-South Educational Research Association, Bowling Green, KY.

- Murphy, J. (2004). *Leadership for literacy: Research-based practice, preK-3*. Thousand Oaks, CA: Corwin.
- National Institute of Child Health and Human Development. (1996). Thirty years of NICHD research: What we now know about how children learn to read. *Effective School Practices, 15*, 33–46.
- National Reading Panel. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. Washington, DC: National Institute of Child Health and Human Development.
- O'Brien, D. M., & Ware, A. M. (2002). Implementing research-based reading programs in the Fort Worth Independent school district. *Journal of Education for Students Placed At Risk, 7*, 167–195.
- Oregon Reading First Center. (2004). *Review of Open Court*. Eugene, OR: Author. Retrieved from http://oregonreadingfirst.uoregon.edu/downloads/instruction/curriculum_review/core_review/oc2002_core_review.pdf
- Pflaum, S., Walberg, H. J., Karigianes, M. L., & Rasher, S. P. (1980). Reading instruction: A quantitative analysis. *Educational Researcher, 9*, 12–18.
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*. Thousand Oaks, CA: Sage.
- Riepl, J. H., Marchand-Martella, N. E., & Martella, R. C. (2008). The effects of Reading Mastery Plus on the beginning reading skills of students with intellectual and developmental disabilities. *Journal of Direct Instruction, 8*, 29–39.
- Rosenshine, B. (2002). Helping students from low-income homes read at grade level. *Journal of Education for Students Placed At Risk, 7*, 273–283.
- Ross, S. M., Nunnery, J. A., Goldfeder, E., McDonald, A., Rachor, R., Hornbeck, M., et al. (2004). Using school reform models to improve reading achievement: A longitudinal study of Direct Instruction and Success for All in an urban district. *Journal of Education for Students Placed At Risk, 9*, 357–388.
- Scarlato, M. C., & Ashara, E. (2004). Effects of Corrective Reading in a residential treatment facility for adjudicated youth. *Journal of Direct Instruction, 4*, 211–217.
- Sewell, W. H., Haller, A. O., & Ohlendorf, G. W. (1970). The educational and early occupational status attainment process: Replication and revision. *American Sociological Review, 35*, 1014–1027.
- Shippen, M. E., Houchins, D. E., Steventon, C., & Sartor, D. (2005). A comparison of two Direct Instruction reading programs for urban middle school students. *Remedial and Special Education, 26*, 175–182.
- Singer, J. D. (1998). Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. *Journal of Educational and Behavioral Statistics, 23*, 323–355.
- Smith, S., Simmons, D., Gleason, M., Kameenui, E., Baker, S., Sprick, M., et al. (2001). An analysis of phonological awareness instruction in four kindergarten basal reading programs. *Reading and Writing Quarterly, 17*, 25–50.
- Snider, V. E. (1990). Direct instruction reading with average first graders. *Reading Improvement, 27*, 143–148.
- Snow, C. E., Burns, M. S., & Griffin, P. (Eds.). (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.
- Stanovich, K. E. (1994). Romance and reality. *Reading Teacher, 47*, 280–291.
- Stebbins, L. B., St. Pierre, R. G., Proper, E. C., Anderson, R. B., & Cerva, T. R. (1977). *Education as experimentation: A planned variation model* (Vol. 4-A). Cambridge, MA: Abt Associates.
- Steventon, C. E., & Fredrick, L. D. (2003). The effects of repeated readings on student performance in the Corrective Reading program. *Journal of Direct Instruction, 3*, 17–27.
- Stockard, J. (2009). *Direct Instruction and first grade reading achievement: The role of technical support and time of implementation*. Eugene, OR: National Institute for Direct Instruction.
- Stringfield, S. C., & Yakimowski-Srebnick, M. (2005). Promise, progress, problems, and paradoxes of three phases of accountability: A longitudinal case study of the Baltimore city public schools. *American Educational Research Journal, 42*, 43–75.
- Strong, A. C., Webby, J. H., Falk, K. B., & Lane, K. L. (2004). The impact of a structured reading curriculum and repeated reading on the performance of junior high students with emotional and behavioral disorders. *School Psychology Review, 33*, 561–581.
- Vitale, M. R., & Joseph, B. L. (2008). Broadening the institutional value of Direct Instruction implemented in a low-SES elementary school: Implications for scale-up and reform. *Journal of Direct Instruction, 8*, 1–18.
- Wolf, F. M. (1986). *Meta-analysis: Quantitative methods for research synthesis*. Newbury Park, CA: Sage.

APPENDIX A: METHODOLOGICAL APPENDIX

The dataset available for analysis had information on NCE and raw scale CTBS scores for first graders and fifth graders in the BCPSS for 1998 through 2003, the schools that the students attended in each year, and a separate file with school characteristics. No information was provided for grades two through four, nor, as noted in the text, was individual socio-demographic information. It was possible, however, to link identification numbers for students between the years, and that was the basis for forming the dataset used for analysis in the article. From this linking process, it was possible to determine which children were in BCPSS in 1998 or 1999 and again in 2002 or 2003. It was also possible to determine which children made expected progress (in first grade in 1998 and in fifth grade in 2002 or in first grade in 1999 and in fifth grade in 2003) and which children had been in first grade in 1998 but were not in fifth grade until 2003, thus apparently repeating a grade in the interim. Given the limitations of the dataset, it was not possible to determine which of the students in first grade in 1998 had previously repeated the grade, nor was it possible to determine which of the students in first grade in 1999 repeated a grade that year. Thus, our information on students who were retained was limited.

Table A1 gives the number of students in the total study population by year and group. The left columns of the table give the raw numbers and the right columns give the percentages. As explained in the text, students in the control group who were in first grade in 1998 (the first three rows of data) had a mixed curriculum in first grade, whereas those in first grade in 1999 had Open Court. The last two rows of the table give the number of students who could be assumed to have been retained. The figures were slightly smaller in the DI schools. The percentages in Table A1 use the total number of students (starting first grade in 1998 or 1999) in each group

TABLE A1
Population and Sample Size by Year and Curriculum

	<i>First Grade Curriculum</i>			
	<i>N</i>		<i>%</i>	
	<i>DI</i>	<i>Control</i>	<i>DI</i>	<i>Control</i>
First grade 1998, no fifth grade data	283	3,641	16.6	23.2
First grade 1998, in BCPSS in fifth grade, different school	151	1,840	8.9	11.7
First grade 1998, in BCPSS in fifth grade, same school	256	2,197	15.0	14.0
First grade 1999, no fifth grade data	542	4,168	31.8	26.6
First grade 1999, in BCPSS in fifth grade, different school	174	1,303	10.2	8.3
First grade 1999, in BCPSS in fifth grade, same school	244	1,875	14.3	12.0
Retained (First in 1998 and fifth in 2003), different schools	45	544	2.6	3.5
Retained (First in 1998 and fifth in 2003), same schools	8	116	0.5	0.7
Totals	1,703	15,684	100.0	100.0

Note. DI = direct instruction; BCPSS = Baltimore City Public School System. These figures omit students who were in schools that began DI after they were in first grade ($n = 341$). Control students in first grade in 1998 had a mixed reading curriculum; control students in first grade in 1999 had Open Court.

as a base. When the number of first grade students in 1998 is used as a base (743 for the DI schools and 8,338 for the control schools), the results indicate the following: For students who were in first grade in DI schools, 6.1% of those who attended a different BCPSS school in fifth grade and 1.1% of those who were in the same school in fifth grade were retained. For students who were in first grade in the control schools, 6.5% of those who attended a different BCPSS school in fifth grade and 1.4% of those who were in the same school in fifth grade

TABLE A2
Scale Scores and NCE Scores by Grade and Group

	Scale Scores						NCE Scores					
	DI			Control			DI			Control		
	<i>M</i>	<i>SD</i>	<i>Med</i>	<i>M</i>	<i>SD</i>	<i>Med</i>	<i>M</i>	<i>SD</i>	<i>Med</i>	<i>M</i>	<i>SD</i>	<i>Med</i>
First grade reading vocabulary												
First 1998, left BCPSS	518	69	516	529	71	524	30	21	27	34	23	31
First 1998, transferred	536	65	538	554	69	552	36	21	36	42	22	41
First 1998, same school	547	68	545	571	65	572	39	22	38	47	22	48
First 1999, left BCPSS	533	65	531	542	73	538	34	21	32	37	23	35
First 1999, transferred	575	60	572	579	68	586	48	20	47	49	22	52
First 1999, same school	577	61	579	595	61	594	48	21	49	54	21	54
Not retained	544	68	545	555	72	552	38	22	37	42	24	41
Retained	514	49	516	514	59	516	27	16	27	28	18	27
First grade reading comprehension												
First 1998, left BCPSS	496	95	504	503	98	515	29	19	24	31	20	28
First 1998, transferred	515	99	525	535	89	544	34	19	31	38	20	38
First 1998, same school	541	80	544	557	80	561	39	20	38	43	20	44
First 1999, left BCPSS	515	90	525	513	100	525	32	19	30	32	21	30
First 1999, transferred	562	73	569	556	85	569	44	18	46	43	20	46
First 1999, same school	569	78	578	574	80	578	46	19	48	47	20	48
Not retained	529	90	544	531	95	544	36	20	36	37	21	36
Retained	485	80	504	488	88	504	25	15	24	26	16	24
Fifth grade reading vocabulary												
First 1998, transferred	634	36	636	633	35	636	44	18	43	43	17	43
First 1998, same school	646	40	644	644	35	644	50	21	48	49	18	48
First 1999, transferred	640	35	644	640	35	644	47	17	48	47	17	48
First 1999, same school	654	36	652	647	32	652	54	19	53	50	17	53
Not retained	645	38	644	642	35	644	49	18	48	48	17	48
Retained	608	54	627	632	36	636	33	20	39	42	17	43
Fifth grade reading comprehension												
First 1998, transferred	640	38	642	641	38	645	44	19	44	45	18	45
First 1998, same school	654	38	657	651	34	651	51	19	52	49	17	49
First 1999, transferred	646	32	648	643	36	645	47	16	47	45	17	45
First 1999, same school	656	36	657	652	31	654	52	17	52	50	16	50
Not retained	650	37	654	647	35	651	49	18	50	48	17	49
Retained	632	25	627	626	34	630	39	13	36	37	16	37

Note. The “not retained” group includes all other students within the system but, as noted in the text, may include children who were retained in earlier years. *Ns* for each sub-group are given in Table A-1. NCE = Normal Curve Equivalent; DI = Direct Instruction; BCPSS = Baltimore City Public School System.

TABLE A3
Average Percentage of Children Receiving Free or Reduced Lunch by Group

	<i>DI</i>			<i>Control</i>		
	<i>M</i>	<i>SD</i>	<i>Med</i>	<i>M</i>	<i>SD</i>	<i>Med</i>
First 1998, left BCPSS	78.2	13.6	82.2	74.7	14.0	76.7
First 1998, transferred	80.7	9.7	84.4	77.5	10.3	78.5
First 1998, same school	75.5	17.0	82.2	71.0	16.1	74.7
First 1999, left BCPSS	80.9	10.3	82.8	74.5	14.1	76.7
First 1999, transferred	81.8	9.7	84.4	76.5	11.1	77.8
First 1999, same school	74.8	17.9	82.2	71.0	15.4	74.7
Not retained	77.5	14.3	82.2	73.9	14.0	77.8
Retained	83.7	4.3	83.7	70.6	16.0	73.9
Total	78.8	13.5	82.8	74.1	14.1	76.3

Note. DI = Direct Instruction; BCPSS = Baltimore City Public School System.

were retained. (These percentages are based on the assumption that students who were in first grade in 1998 and in fifth grade in 2003, six years later, had likely repeated a grade in the interim.)

Table A2 reports the average raw scale and NCE scores for the CTBS reading comprehension and vocabulary tests in first and fifth grade for students in both the DI and control schools. Results are disaggregated by year and by whether or not a child remained in BCPSS and in the same school in the first and fifth grade samples. Thus, the data in this table allow one to compare scores of students used in our analysis (those who remained in the same school throughout the elementary career in BCPSS) and other students. As explained in the text, students who remained in the same schools from first to fifth grade had higher achievement scores than other students.

The bottom two lines of each section of Table A2 compare the scores of students who were determined to have been retained and the average of all other students. Because the sample size of retained students was so small, scores of those who had transferred to other schools and those who remained in the same schools were combined for this analysis. In all comparisons, the scores of the students who were retained were lower than those of students who were not. As noted in the text, the students who were retained were not included in the analyses.

Finally, Table A3 reports the average percentage of students receiving FRL in the schools attended by group.