

Technical Report #2007-1:
Academic Acceleration in First Grade Reading Using the
Direct Instruction Model

Michael Rebar
Eastern Washington University

The preparation of this manuscript was supported in part by a grant (Award No. S332B050017) to the Center for Applied Research in Education (CARE), Hillsboro, Oregon from the Office of Elementary and Secondary Education (OESE), U.S. Department of Education.

**Academic Acceleration in First Grade Reading Using the
Direct Instruction Model (Technical Report #2007-1)**

Current Study

This study compares the first grade reading outcomes of school reform efforts in Baltimore from the 1996-1997 school year to the 2002-2003 school year for eleven schools that implemented the Direct Instruction model developed by the National Institute for Direct Instruction (NIFDI), the successor to the Direct Instruction model provider in Project Follow Through. The eleven treatment schools were managed by the Baltimore Curriculum Project, which contracted with the NIFDI to provide preservice and inservice training, coaching, and technical assistance.

This study provides a quasi-experimental comparison of the performance of eleven NIFDI schools (the *NIFDI* condition) with two comparison groups: (a) eleven other Baltimore schools matched on achievement, poverty and ethnicity (the *Matched Control* condition), and (b) all Baltimore elementary schools with the exception of the NIFDI schools (the *Baltimore* condition). Note that the *Matched Control* schools are included in the *Baltimore* condition; they represent a significant proportion of the low-achieving schools in the district and their exclusion from the *Baltimore* condition would artificially inflate apparent district achievement. The *Baltimore* condition is important because during the same period, Baltimore Public Schools adopted and implemented *Open Court*, an empirically validated program.

Previous Studies

Two earlier studies have reported on school reform in Baltimore during the same time period, 1996 through 2002 (Mac Iver et al., 2003; Addison & Yakimowski, 2003). This study differs from that of Mac Iver's in three ways: (a) the Mac Iver study included six schools, rather than the full eleven NIFDI schools; (b) the Mac Iver study compared two small cohorts spanning different grade ranges (k-2 and 3-5) -- an inappropriate comparison; and (c) the Mac Iver study focused almost exclusively on results attained by 2000 and terminated in 2002. In addition, the current study differs from the Addison and Yakimowski report in three key ways: (a) their report erroneously includes 16 schools in the Direct Instruction condition, and (b) their report is based on a series of sub-analyses without providing a global analysis, and (c) their report ends with the 2001-2002 school year. Note that there were 16 Direct Instruction schools, but only eleven were supported by NIFDI. Five schools were supported by two other providers.

Research Questions

The current study addresses two questions about implementation with regard to beginning reading:

1. How do the effects of the Direct Instruction model on first grade reading performance compare with the effects of other reform models?
2. How long does it take to reach grade-level performance with the NIFDI model in very low-achieving schools?

Project Overview

Throughout the early 1990s, student achievement on high stakes achievement tests was chronically low throughout the Baltimore City Public School System (BCPSS). The majority of BCPSS students performed substantially below grade level in reading and math (Lambert & Reynolds, 1997). Baltimore city schools had markedly lower achievement than other districts in the state. The Abell Foundation created the Baltimore Curriculum Project (BCP) to develop and test an approach to school reform that could later be scaled and used in Baltimore city schools. The BCP selected Direct Instruction as the core curriculum and NIFDI to provide training and monitoring of DI implementation.

Method

Participants

School Sites. Eleven urban elementary schools in Baltimore constituted the treatment schools (*NIFDI* condition). Eleven other elementary schools individually matched on poverty and ethnicity characteristics and on initial achievement served as control schools (*Matched Control* condition). All 122 elementary schools except those in the NIFDI condition served as a secondary comparison group (*Baltimore* condition). Note that treatment was phased in over three years. In 1997 there were four schools in each condition. In 1998 there were six schools in each condition. From 1999 on there were eleven schools in each condition.

Student Participants. More than 41,223 students enrolled between 1997-1998 and 2002-2003 served as participants. The number of participants by condition is summarized in Table 1. As the table shows, the number of students varied each year by condition. In general, NIFDI schools had a lower enrollment of first grade students. Three of the *Matched Control* schools were substantially larger than the treatment schools. The total number of first graders in the treatment group across years was 2,824 with an annual mean of 471 (SD = 97.59). The total number of first graders in the *Matched Control* condition across years was 3,922 with an annual mean of 654 (SD = 121.72). Recall that the *Matched Control* and *Baltimore* conditions overlap, so the columnar totals are the sum of participants in the *NIFDI* and *Baltimore* conditions.

Table 1.

Total Number of Students by Year and Condition (Grade 1)

Condition	Year						Total
	1998	1999	2000	2001	2002	2003	
NIFDI	289	570	473	535	494	463	2,824
Matched Control	421	698	740	707	621	735	3,922
Baltimore	7,125	7,056	6,633	6,262	5,734	5,589	38,399
Total	7,414	7,626	7,106	6,797	6,228	6,052	41,223

The *NIFDI* and *Matched Control* condition schools had similar Free and Reduced Lunch rates, an indicator of socioeconomic status (SES) at the

beginning and *end* of the project. In the middle of the study, however, NIFDI schools had a substantially higher Free and Reduced Lunch rate for three years (1998-1999 through 2001-2002).

Table 2.

Percentage of Free and Reduced Lunch Rate by Year and Condition

Condition	Year					
	1998	1999	2000	2001	2002	2003
NIFDI	81	84	81	81	88	85
Control	79	71	75	73	85	83
Other Baltimore	72	72	72	71	75	74

NIFDI schools self-selected Direct Instruction as their reform model. The NIFDI group is an exhaustive sample of all eleven schools that implemented the Direct Instruction model according to NIFDI guidelines. Five other schools selected the NIFDI model, but did not continue with NIFDI through the entire treatment period. Those schools are excluded from the *NIFDI* condition and are included in the *Baltimore* condition.

A summary of ethnicity by condition is provided in Table 3. For the current study, the ethnicity of participants was coded as (a) White, (b) African American, (c) Hispanic, or (d) Other. The district identified two additional categories: Asian and Native American. Because of their low incidence, those were collapsed into

the Other category. Students in the Baltimore Public Schools are predominantly African American.

Table 3.

Ethnicity of Participants (Percentage) by Year and Condition

Condition		Year					
		1998	1999	2000	2001	2002	2003
NIFDI	White	21	20	17	16	15	13
	African American	73	74	75	76	77	78
	Hispanic	3	4	4	5	5	6
	Other	3	2	4	3	3	3
Matched Control	White	17	17	14	13	12	11
	African American	81	81	82	84	85	86
	Hispanic	1	1	2	2	2	3
	Other	1	1	2	1	1	0
Baltimore	White	15	13	12	12	12	11
	African American	83	85	85	86	86	87
	Hispanic	0	0	1	1	1	1
	Other	2	2	2	1	1	1

The *Baltimore* condition comprised the remaining Baltimore Public School System elementary schools; this number ranged from 103 to 122 schools during the study.

Measures

No measurement plan was in effect at the start of the project. During the course of the project, three measures of reading were employed: (a) the *Peabody Picture Vocabulary Test*, (b) the fourth edition of the *Comprehensive Test of Basic Skills*, and (c) the fifth edition of the *Comprehensive Test of Basic Skills* also known as the *TerraNova*. All of these measures were administered in the spring of each year.

In the spring of 1997, the Peabody Picture Vocabulary Test (PPVT) was administered to kindergartners in the *NIFDI* and *Matched Control* conditions. The PPVT is a widely used norm-referenced picture identification test that is highly predictive of future reading ability. No other norm-referenced test was administered in the 1996-1997 school year. No tests were administered in the Fall of 1997. The results serve as the best estimate of initial achievement level (i.e., prior to intervention).

In the spring of 1998 and 1999 the *Comprehensive Test of Basic Skills, Fourth Edition (CTBS/4)* (CTB, 1991) was administered to all Baltimore first graders. The *Comprehensive Test of Basic Skills, Fifth Edition (CTBS/5–TerraNova)* was administered to all first graders in 2000 through 2003.

Table 4.

Distribution of Measures Across Time (Spring Administrations)

Measure	Year						
	1997	1998	1999	2000	2001	2002	2003
PPVT	X						
CTBS/4		X	X				
CTBS/5 TerraNova				X	X	X	X

Note: The PPVT was administered to kindergartners in spring of 1997 who were first graders in 1997-1998.

Intervention

Treatment. Students in the treatment condition received two 30-minute periods of Direct Instruction reading per day in kindergarten and first grade throughout the study. The periods were not consecutive. Schedules varied by site.

Treatment was phased in over a 3-year period. Of the eleven treatment schools, four began treatment in 1996-1997, two in 1997-1998, and five in 1998-1999. See Table 5 below (p. 13) for details on treatment onset; a blank cell indicates that a school was not yet part of the treatment condition.

All treatment schools used the *Reading Mastery Classic* curriculum (SRA McGraw-Hill), a Direct Instruction reading program, for all project years. *Reading Mastery Classic* is a scripted, mastery-based core reading program that focuses on decoding and comprehension. It utilizes a specialized orthography to help students discriminate between otherwise confusing letters and letter

combinations. 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 other 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 high rate of active responding. Student responses during independent seat work are also closely monitored and immediately remediated.

The students in this study also received language instruction for 30 minutes per day in *Language for Learning* and *Language for Thinking* (published by SRA). Both *Language for Learning* and *Language for Thinking* are general knowledge programs that focus on oral language development.

Reform Model. When a school implements the NIFDI model, instructional programs are phased in over several years. In year one, language and reading programs are introduced in kindergarten and first grade.

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 inservice sessions that address problems teachers are experiencing. There also are weekly

conference calls that address all the current problems each classroom is experiencing in meeting projected performance gains.

During the second year, teachers who perform well are identified as coaches and are deployed to work with other teachers in the school. In addition, the rest of the Direct Instruction curricula are implemented (e.g., math and a program on general information).

NIFDI attempts to place the school's best teachers in kindergarten and first grade because these grades are assumed to be the most important for achieving continued acceleration of student performance. NIFDI does not believe that it is possible to fully implement curriculum reform in the early grades of a failed school in less than 3 to 5 years.

All teachers receive preservice training and coaching until they teach each program to a minimum adequate level of fidelity. Teachers continue to receive inservice coaching to improve implementation fidelity.

Instruction for Comparison Conditions. Prior to 1998, the *Matched Control* and *Baltimore* condition schools were free to use any curriculum program desired. There was no district-wide structured reading program and schools used a variety of instructional programs. In the fall of 1998, the district adopted *Open Court Reading* in kindergarten through second grade. The 2002 edition of *Open Court* was favorably reviewed by the Florida Center for Reading Research (FCRR, 2004) and the Oregon Reading First Center as a core instructional program for Reading First (Oregon Reading First Center, 2004).

Results

Analyses

Descriptive statistics including the mean, median, standard deviation, and count were computed for each of the eleven treatment schools and are reported in Appendix A. The mean, median, and standard deviation values for *Total Reading* on the *CTBS* are expressed as normal curve equivalent (NCE) scores. NCE values are reported because they represent equal interval data that can be properly used in the calculation of means and standard deviations. Note that five schools (Schools 4, 5, 6, 10 and 11 in Table A-1, p. 29) did not begin treatment until the 1998-1999 school year; data for these schools are not reported for the 1997-1998 school year.

Summarized descriptive statistics for each condition are reported in Table 5 below. Disaggregated statistics by school are not reported.

Table 5.

Descriptive Statistics on CTBS Total Reading Scores (NCE) by Condition Across Time

Condition		Year					
		1998	1999	2000	2001	2002	2003
NIFDI	Mean	30.31	37.17	48.62	56.68	62.87	64.03
	Median	28	37	48	57	63	64
	SD	18.27	19.01	21.75	21.13	23.34	23.72
	Count	289	570	473	535	494	463
Matched Control	Mean	29.05	33.69	41.52	42.72	48.26	45.30
	Median	27	32	41.50	42	48	45
	SD	17.84	20.71	21.39	21.93	21.47	22.43
	Count	421	698	740	707	621	735
Baltimore	Mean	36.94	40.90	48.77	52.16	53.80	55.14
	Median	35	42	49	53	55	55
	SD	21.00	21.60	22.69	22.38	22.68	22.25
	Count	7,125	7,056	6,633	6,262	5,734	5,589

A series of two-tailed independent-samples *t* tests ($\alpha=.05$) were used to test for differences in mean performance between the *NIFDI* and *Matched*

Control conditions for each year. The CTBS Total Reading NCE score was used in all *t* tests.

Table 6.

Test of Mean Differences on CTBS Total Reading Scores (NCE) Between NIFDI and Matched Control

Year	<i>NIFDI</i>			<i>Matched Control</i>			df	t	Sig.
	N	Mean	SD	N	Mean	SD			
1998	289	30.31	18.27	421	29.05	17.84	708	.913	.362
1999	570	37.17	19.01	698	33.69	20.71	1266	3.084	.002*
2000	473	48.62	21.75	740	41.52	21.39	1211	5.601	.000*
2001	535	56.68	21.13	707	42.72	21.93	1240	11.283	.000*
2002	494	62.87	23.34	621	48.26	21.47	1113	10.86	.000*
2003	463	64.03	23.72	735	45.30	22.43	1196	13.77	.000*

Note: * Statistically significant at the 0.01 level after taking multiple comparisons into account.

Effect sizes were estimated using an extension of *Glass' delta*, a commonly accepted metric that represents effect sizes in standard deviation units of a *control group*, in this case *the pooled group of all NIFDI and Matched Control condition students in 1997-1998*. Therefore, *Glass' delta* was calculated by subtracting the mean of each condition each year from mean of the pooled conditions for 1998 and dividing the result by the pooled standard deviation for 1998. The pooled mean serves as the subtrahend in the numerator of each

calculation. The pooled standard deviation serves as the fixed denominator for each calculation.

The underlying logic in using the pooled mean and standard deviation from 1998 for all calculations is to fix the baseline for comparisons within years and across years. This extension of *Glass' delta* was selected because there is no single accepted technique for estimating effect sizes for systemic school reform with varied groups over time. It is important to note that the effect size estimates using *Cohen's d* or other estimates, would not be significantly different.

Table 7.

Cumulative Effect Sizes (Glass' delta) by Year and Condition

Condition	Year					
	1998	1999	2000	2001	2002	2003
NIFDI	.04	.42	1.06	1.5	1.85	1.91
Matched Control	-.03	.23	.66	.73	1.04	.87

Table 8.

Cumulative Effect Sizes (Glass' delta) by Year and Condition Adjusted for Common Intercept (Zero Point)

Condition	Year					
	1998	1999	2000	2001	2002	2003
NIFDI	0.0	.38	1.02	1.46	1.81	1.87
Matched Control	0.0	.26	.69	.76	1.07	.90
Difference	0.0	.12	.33	.70	.74	.97

All effect sizes are referenced to the baseline performance on the CTBS/4 in the spring of 1998. The effect size the *Matched Control* schools achieved is more than three-fourths of a standard deviation. The effect of the NIFDI intervention was large, nearly two standard deviations.

Summary. Research question one asked “How do the effects of the Direct Instruction Model on first grade reading performance compare with the effects of other reform models?” The NIFDI model achieved significantly higher scores within one year and maintained over the six remaining years of the study. The difference in the performance of students in the NIFDI model was not only statistically significant in all years but also educationally significant. According to Wolf (1986), educational significance is achieved when an effect size is greater than one quarter standard deviation. In the first year after implementation, the difference between the gains achieved by the *NIFDI* and *Matched Control*

schools was only .12 standard deviations. In all subsequent years, the difference increased to nearly one full standard deviation by year 6. The average student performance level in the *NIFDI* schools moved from the lowest quintile to well above the national mean. All *NIFDI* schools showed substantial achievement gains and several became beacon schools for the district. The three highest achieving schools in the district in 2003 were *NIFDI* schools. The *Matched Control* schools never achieved a mean score above the national average.

In all, 49 of the 103 *Baltimore* condition schools (48%) achieved a mean *Total Reading* NCE score of 55 or greater in the 2003 school year. In the *NIFDI* condition, ten of eleven schools (or 91% of these schools) reached this level. Only one of the eleven *Matched Control* condition schools reached an NCE score of 55; this school ranked 49th in the *Baltimore* condition. A one-tailed Fisher's Exact Test, a non-parametric statistical test, was used to test the null hypothesis that there were no differences in the percentage of schools in each condition that reached a CTBS Total Reading NCE score of 55 or greater, a score at which the confidence interval would be entirely above the mean. The exact probabilities are presented below in Table 9.

Table 9.

Fisher's Exact Test for Probability of Same Likelihood of Exceeding Mean by Condition

Comparison	Probability
NIFDI, Matched Control	0.0010
NIFDI, Baltimore	0.0304

The Fisher's Exact Test indicates that schools in the NIFDI condition were significantly more likely to exceed the national mean than either the *Baltimore* or *Matched Condition* schools. *The three highest achieving elementary schools in the district were NIFDI schools, achieving mean Total Reading NCE scores of 92, 90, and 83.* The highest three schools in the *Baltimore* condition achieved scores of 78, 76, and 75. It is perhaps ironic that the highest achieving *Baltimore* school (with a mean NCE score of 78) is one of the DI schools that did not continue as a NIFDI school but continued to use the Direct Instruction reading program.

Research question two asked "How long does it take to reach grade-level performance with the NIFDI model in very low-achieving schools?" If *reach grade-level performance* is operationalized as having the sample mean exceed the national mean, the NIFDI model achieved this goal within two to four years. In contrast, the *Matched Control* schools approached the mean in the fifth year of implementation, but never exceeded it. The *Baltimore* condition schools took four years to achieve mean performance, even though their students started with

higher achievement levels and they began their reforms in 1996-1997. In addition, these schools show a level achievement trend from 2000-2001 to 2002-2003.

Discussion

Measuring Reform

Comparison Schools. During this study, schools in the *Baltimore* condition improved their first grade *Total Reading CTBS* NCE scores from 36.94 to 55.15 representing growth from the 24th percentile in 1997-1998 to the 59th percentile in 2002-2003, a remarkable achievement for a large district. The trend in achievement appears to flatten substantially in 2000-2001 through 2002-2003, as the district reaches the fiftieth percentile. Complete percentile scores are reported in Table 10 below.

The eleven schools in the *Matched Control* condition showed good growth in achievement, increasing their *Total Reading CTBS* NCE scores from 29.05 to 45.30, representing growth from the 14th to 40th percentile. As appears evident with the *Baltimore* condition schools, the *Matched Control* condition schools show a decreased growth rate as they near the national mean in 2001, and a decline in the 2002-2003 school year.

NIFDI Schools. Grade one performance in the NIFDI schools showed a increase in Total Reading NCE score from 30.31 to 64.03 (an increase from the 15th percentile in 1997 to the 75th percentile in 2003), roughly sixty percentile points in seven years. The growth in achievement scores for NIFDI schools did

not level off as students reached the national mean, but continued to grow through the end of the project period, at which point they were approximately one-half standard deviation above the national mean and nearly a full standard deviation above the *Matched Control* schools. The increasing effect size each year shows that implementation improved from year to year. Strong gains were evident each year through 2001-2002, which was the final year of NIFDI's participation with all schools. In 2002-2003, scores increased approximately two percentile points (see Table 10).

Limitations. A significant limitation of this study is the reliance upon a single measure of reading achievement for estimating the impact of a complex intervention such as the Direct Instruction full-immersion model developed by NIFDI. Further studies should utilize reading progress monitoring measures and report on in-program test data.

Even though the reading achievement measures changed over time, cautious comparisons can be made. Table 10 below extends the comparison of CTBS scores presented in Table 5 to include the PPVT scores for 1996-1997. For ease of interpretation, the mean NCE scores for each condition were converted to percentile scores.

Table 10.

*Median Grade One CTBS Total Reading Percentile Scores on Reading**Measures by Year and Condition*

Condition	Year						
	1997	1998	1999	2000	2001	2002	2003
NIFDI	15*	15	27	46	63	73	75
Matched Control	26*	14	20	34	35	46	40
Baltimore	27*	24	35	48	56	59	59

*Note: The PPVT was administered in 1996-1997, the CTBS/4 in 1998-1999 and 1999-2000 and the CTBS/5 TerraNova thereafter.

Similarly, NIFDI observed that there was a strong correlation between student achievement and the extent to which the schools implemented the NIFDI model. The schools with the more thorough implementations went farther through the reading and language programs and achieved higher scores. Even with considerable variation in degree of implementation, the NIFDI full immersion model had a strong positive impact on early reading and accelerated achievement compared to another effective program, *Open Court*. The inclusion of fidelity measures in future studies would help clarify the relative contributions of components of the NIFDI model. Future studies should also use specific measures of implementation fidelity.

Scaling Reform: Lessons from Empirical Evidence

While the results of this study provide evidence that school reform can be achieved and sustained with the NIFDI model, bringing an academic reform to scale poses additional challenges. Despite the gains demonstrated with the NIFDI model in Baltimore in 1997 through 2003, the district did not systematically expand the model to other schools. In fact, during the course of this study the district mandated procedures that were at odds with model provisions. Teachers were required to attend district inservice trainings that advised them to do things differently than what the NIFDI model specifies. The central administration required schools to provide daily test-preparation periods for the entire school year. NIFDI advised the principals in their model not to do this because it believed greater gains would be possible by implementing the model rather than

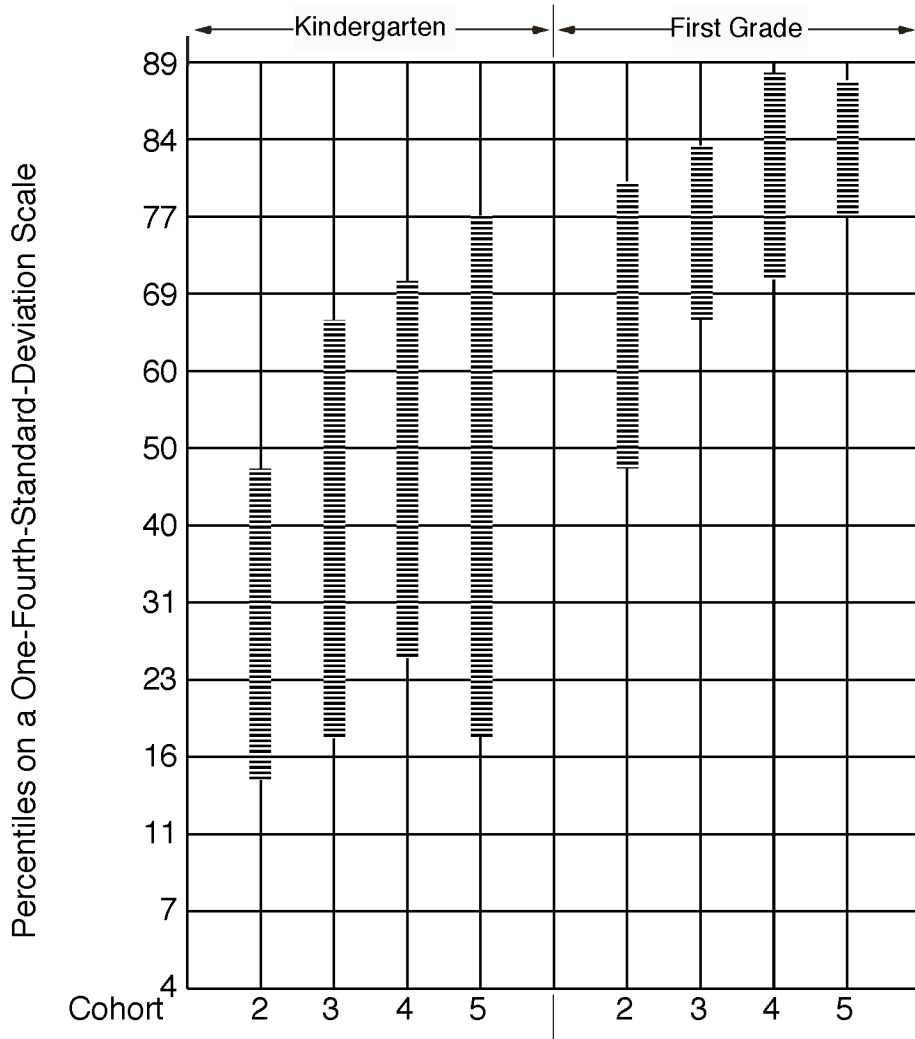
providing test preparation (which would also tend to artificially inflate achievement scores). In fact, the six principals who implemented the model most faithfully and ignored some of the district mandates achieved better performance than any of those who complied with district regulations.

District-level barriers also occurred in Project Follow Through, the largest educational study ever conducted. It compared the performance of the Direct Instruction model with that of 21 other models in a study of how to accelerate performance of at-risk children in kindergarten through third grade. Becker and Engelmann (1978) concluded that none of the 20 districts that adopted the DI model ever fully implemented it because of district interference, even though the model sites were supposed to have received waivers on conflicting practices.

In spite of the barriers to implementation, Follow Through showed that reform can be effectively attained, scaled, and sustained across diverse sites using the Direct Instruction model. Figure 1 below, excerpted from Technical Report 78-1 (Becker & Engelmann, 1978, p. 75) illustrates student achievement increases in kindergarten and first grade reading performance over a two-year period for four cohorts (numbered 2-5). Note that the y-axis in Figure 1 uses a one-fourth standard deviation unit scale, an equal interval scale, so that percentile scores can be displayed and visually interpreted without distortion. A plot using NCE scores would be almost visually identical, but NCE scores are less familiar than percentiles.

Figure 1.

Norm-referenced gains on WRAT Reading by cohort from kindergarten through first grade for K-starting, low-income, DI Follow Through students.



The trend is similar to that reported in the current study. Substantial improvement occurred every year. Furthermore, the Direct Instruction model was scaled to a greater degree in the Follow Through study; the cohort analysis

includes schools in thirteen different districts in ten different states. The number of subjects in each cohort is summarized in Table 11 below.

Table 11.

Total Number of Students in Each Cohort by Grade (DI Follow Through Model)

Grade	Cohort			
	2	3	4	5
K	755	1184	1173	927
1	758	1135	988	889

Each cohort made remarkable gains, starting kindergarten in the lowest quartile and ending first grade in the highest quartile. Summarized percentile scores for each cohort are presented in Table 12 below.

Table 12.

WRAT Reading Percentile Scores for Each Cohort (DI Follow Through Model)

Grade	Cohort			
	2	3	4	5
K (Fall)	14	18	25	18
1 (Spring)	79	82	88	87

The absolute performance level is not directly comparable to that of the current study, however, because the Follow Through trend is based upon the *Wide Range Achievement Test (WRAT)*, a measure of oral-reading accuracy. Note also that the DI Follow Through model provided more intensive training and coaching support than NIFDI provided in Baltimore, roughly twice the number of person days of support. The trend for the Follow Through schools and the relative acceleration of NIFDI schools show that performance continues to improve for four years. This trend implies that the teachers in the later years were likely more effective than the same teachers were in the earlier years. The details of why and how this occurs require further investigation. NIFDI staff believe that the improvement in performance hinges upon very specific skills and practices that teachers are taught and specific changes in the ethic and working relationships between principal, coordinator, teachers, and NIFDI personnel. These goals are not fully achieved after only one or two years as demonstrated by both the current study and the Follow Through study. The intensive training and support provided in the Follow Through study resulted in substantially greater gains over a two-year period than those achieved in Baltimore over a six-year period. These greater gains may be due to greater consistency in the quality of the implementations in the various schools in Follow Through, since three of the NIFDI Baltimore schools achieved top scores and ranks. Yet as a group, the NIFDI Baltimore schools did not seem to achieve at a level comparable to the DI Follow Through schools.

References

- Addison, K., & Yakimowski, M. (2003). An Evaluation of the Direct Instruction Program: A Report Prepared for the Board of School Commissioners. Division of Research, Evaluation, Assessment, and Accountability, Baltimore City Public School System.
- Becker, W. C., & Engelmann, S. (1978). Analysis of achievement data on six cohorts of low-income children from 20 school districts in the University of Oregon Direct Instruction Follow Through Model (Technical Report #78-1). Eugene, OR: University of Oregon, Office of Education, Follow Through Project.
- Florida Center for Reading Research (2004). Summary Table for FCRR Reports: Comprehensive Core Reading Programs. Retrieved June 2, 2007, from <http://www.fcrr.org/FCRRReports/CReports.aspx?rep=core>
- Lambert, C., & Reynolds, J.E. (1997). The state of Baltimore's schools: Data and information on the status of Baltimore City Public Schools. (Vols. 1-3). Baltimore, MD: Advocates for Children and Youth.
- Mac Iver, M., Kemper, E., & Stringfield, S. (2003). The Baltimore Curriculum Project: Final report of a four-year evaluation study. CRESPAR Report #62. Baltimore, MD and Washington, DC: Center for Research on the Education of Students Placed at Risk.
- Oregon Reading First Center (2004, March). Review of comprehensive programs. Eugene, OR: Author. [Electronic version]. Retrieved June 2, 2007, from http://reading.uoregon.edu/curricula/or_rfc_review_n.php.

Wolf, F. M. (1986). *Meta-analysis: Quantitative Methods for Research Synthesis*.
Beverly Hills, CA: Sage.

Appendix A

Table A-1.

*Demographic Statistics for Individual Schools in 1997-1998 and 2002-2003**

School No.	Condition**	Percent African American	Percent Hispanic	Percent White	Percent Free/Reduced Lunch
1	NIFDI	100 (100)	0 (0)	0 (0)	87 (87)
	MC	99 (100)	0 (0)	0 (0)	71 (85)
2	NIFDI	33 (54)	2 (4)	54 (32)	87 (88)
	MC	63 (74)	3 (6)	31 (18)	87 (75)
3	NIFDI	97 (99)	0 (0)	2 (1)	82 (89)
	MC	98 (98)	0 (0)	2 (2)	83 (81)
4	NIFDI	100 (100)	0 (0)	0 (0)	85 (93)
	MC	100 (100)	0 (0)	0 (0)	84 (90)
5	NIFDI	99 (99)	0 (0)	0 (0)	71 (82)
	MC	82 (81)	0 (1)	17 (17)	73 (81)
6	NIFDI	98 (100)	0 (0)	2 (0)	91 (92)
	MC	99 (100)	0 (0)	0 (0)	92 (85)
7	NIFDI	62 (74)	0 (2)	36 (22)	68 (75)
	MC	50 (66)	4 (16)	39 (14)	75 (82)
8	NIFDI	22 (21)	29 (55)	43 (19)	81 (82)
	MC	98 (100)	0 (0)	0 (0)	74 (89)
9	NIFDI	5 (14)	4 (8)	84 (69)	69 (70)
	MC	14 (29)	0 (7)	84 (61)	58 (74)
10	NIFDI	100 (100)	0 (0)	0 (0)	86 (89)
	MC	100 (100)	0 (0)	0 (0)	83 (76)
11	NIFDI	88 (93)	1 (1)	11 (5)	89 (90)
	MC	89 (97)	0 (0)	11 (3)	92 (92)

* Note: 1997-1998 values are shown without parentheses. 2002-2003 values are in parentheses.

** MC signifies the *Matched Control* condition.

Table A-2.

*Summarized Mean Demographic Statistics for Treatment and Matched Control**Schools in 1997-1998*

Condition	Percent African American	Percent Hispanic	Percent White	Percent Free/Reduced Lunch Eligible	Percent 1997-1998 Students Repeating Grade 1998-1999	1997- 1998 CTBS/4 Total Reading (NCE)
NIFDI	73	3	21	81	4	30.31
Matched Control	81	1	17	79	6	29.05

Table A-3.

*Descriptive Statistics for Individual NIFDI Schools on CTBS Total Reading (NCE)
Across Time*

School No.	Desc. Statistic	Year					
		1998	1999	2000	2001	2002	2003
1	Mean	23	45	50	58	54	60
	Median	21	44	50	58	56	62
	SD	14	21	14	18	20	23
	Count	62	47	54	54	63	51
2	Mean	28	47	48	50	60	65
	Median	24	46	50	53	60	66
	SD	17	18	19	21	21	13
	Count	44	52	50	47	51	48
3	Mean	35	41	58	68	70	90
	Median	31	40	65	68	79	99
	SD	18	19	23	23	27	18
	Count	64	52	43	44	72	53

Table A-3.

*Descriptive Statistics for Individual Treatment Schools on CTBS Total Reading
(NCE) Across Time (Continued)*

School No.	Desc. Statistic	Year					
		1998	1999	2000	2001	2002	2003
4	Mean		28	51	38	65	53
	Median		27	51	39	68	52
	SD		16	17	15	24	18
	Count		56	43	52	49	46
5	Mean		40	41	56	61	61
	Median		41	36	57	60	62
	SD		18	22	15	20	20
	Count		68	66	78	57	61
6	Mean		31	30	62	54	37
	Median		29	28	73	55	35
	SD		20	20	21	18	17
	Count		64	49	27	27	33
7	Mean	29	35	61	65	62	59
	Median	29	36	62	63	60	66
	SD	14	12	22	19	21	24
	Count	51	46	31	37	25	33

Table A-3.

Descriptive Statistics for Individual Treatment Schools on CTBS Total Reading (NCE) Across Time (Continued)

School No.	Desc. Statistic	Year					
		1998	1999	2000	2001	2002	2003
8	Mean	38	36	46	54	45	55
	Median	34	36	47	48	41	55
	SD	20	19	20	27	24	17
	Count	30	38	29	26	26	30
9	Mean	33	37	51	53	64	58
	Median	31	35	53	53	63	62
	SD	22	19	21	19	18	20
	Count	61	86	87	73	53	59
10	Mean		40	49	73	77	92
	Median		43	46	75	81	99
	SD		20	23	16	21	14
	Count		63	48	47	42	23
11	Mean		38	59	54	68	83
	Median		37	55	55	75	84
	SD		22	18	24	26	17
	Count		60	37	54	45	29

Figure A-1.

Mean CTBS Total Reading (NCE) scores over time for NIFDI schools.

