

**Memphis Striving Readers Project
Year 2 Evaluation Report**

by

Research for Better Schools

RMC Research Corporation

December 19, 2008

Revised April 30, 1009

Memphis Striving Readers Project Year 2 Evaluation Report

December 2008

Kelly Feighan
Debra Coffey
Research for Better Schools

Allen Schenk
RMC Research Corporation

Jill Feldman
Research for Better Schools

Acknowledgments

This study is the result of strong partnerships forged between staff members at Memphis City Schools (MCS), the University of Memphis, and Research for Better Schools (RBS) and its associates, RMC and Edvantia. The authors represent a small fraction of individuals who contributed to the study's design, implementation, and analysis. The specific contributions of each author are as follows: Kelly Feighan, for her leadership in conducting and summarizing research related to implementation of the whole-school intervention and logistical coordination of project-wide collection of data; Debra Coffey, for her leadership in conducting and summarizing research related to implementation of the targeted intervention and skillful editing of this report; Allen Schenk for his intellectual leadership in developing the research design and analytic models used to assess impact for both interventions; and Jill Feldman for her creative and intellectual leadership as principal investigator of the research team.

In particular, the authors wish to thank Dr. Elizabeth Heeren, Principal Investigator of the Memphis Striving Readers Project at MCS, for her invaluable guidance in all aspects of the implementation evaluations. Others at MCS whom we wish to thank include Ric Potts, who provided invaluable guidance throughout the research process, and Rorie Harris, who prepared MCS data sets for analyses and coordinated the administration of the Iowa Test of Basic Skills twice annually. We also wish to thank the Memphis Content Academy (MCLA) literacy coaches, school principals, course instructors, and participating teachers, all of whom provided information that is essential to our study.

The chief architect of MCLA, Robert Cooter, formerly of the University of Memphis and now at Bellarmine University, has demonstrated unwavering support and enthusiasm for the evaluation of the whole-school intervention in his genuine interest to improve adolescent literacy. We also thank others at the University of Memphis, including Helen Perkins, co-Principal Investigator of MSRP, for her insights and feedback about instruments and other materials used to collect implementation data, as well as Theresa Lavell and Kathy Cooter (now also at Bellarmine University) for their ongoing contributions.

We wish to thank field team members from RBS and Edvantia who in Year 2 of this study conducted classroom observations, collected survey data, or assisted with the analyses. Specifically, at Edvantia, we thank Kristine Chadwick, Juan D'Brot, Aaron Butler, Joe Holloway, Nathaniel Hixson, Akisha Jones, and Charity Sheffield. At RBS, we would like to acknowledge Lindsay Wilhelmi, Eric Kucharik, Christina Shipman, Melissa Padula, and Liesa Stamm.

Lastly, the authors of this report wish to acknowledge the time and efforts of our technical working group members who convened monthly to guide the study's analyses. In addition to colleagues listed above who participate in this working group, we wish to thank Keith Kershner of RBS, Jim Craig at the School of Education at Virginia Tech, and Michael Frye of RMC. Cris Price, Barbara Goodson, and Julie Meltzer from Abt, our technical assistance provider, have also been exceedingly helpful in shaping our analyses and ensuring that we produce work that meets high standards of rigor.

Table of Contents

- I. Executive Summary..... 1**

- II. Introduction and Study Background..... 5**
 - Introduction.....5
 - Background Context.....6
 - Theoretical Rationale for and Description of the Intervention Models.....6
 - Description of the Targeted Intervention.....6
 - Students Targeted by the Intervention.....6
 - Logic Model for Targeted Intervention7
 - Professional Development Model Components.....7
 - Classroom Model Components.....8
 - Theoretical Rationale for and Description of Whole-School Intervention Model11
 - The Innovation Configuration Map.....12
 - Professional Development Model Components.....13
 - Classroom Instruction Model Components14
 - Changes Made to the Professional Development Model between Years 1 and 215
 - Brief Overview of Key Evaluation Design Features.....15
 - READ 180* Impact.....15
 - MCLA Impact.....16

- III. Evaluation of the Implementation of the Targeted Intervention, Years 1 and 2..... 17**
 - Summary of the Design17
 - Development of the Ratings and Scale for Years 1 and 2.....18
 - Year 1 Implementation Study.....20
 - Professional Development Levels20
 - Levels and Variability of Implementation at the Classroom Level20
 - Year 2 Implementation Study.....23
 - Professional Development Levels23
 - Levels and Variability of Implementation at the Classroom Level23
 - Conclusions Regarding Implementation of the Targeted Intervention25
 - Description of the Counterfactual for the Targeted Intervention.....26

IV. Evaluation of the Impacts of the Targeted Intervention: Years 1 and 2.....	28
Study Design.....	28
Sample Selection	28
Data Collection	30
Data Analysis	31
Selection of Covariates	33
Treatment of Missing Data	34
Description of the Year 1 and Year 2 Samples of Students for <i>READ 180</i> Impact Analyses	34
Equivalence on Student Demographic Characteristics	34
Equivalence on Baseline Achievement	36
Impact of <i>READ 180</i> Participation on Student Achievement in Years 1 and 2.....	38
Immediate Impact of <i>READ 180</i> in Year 1	38
Immediate Impact of <i>READ 180</i> at Each Grade in Year 1	40
Immediate Impact of <i>READ 180</i> at Sixth Grade in Year 2	41
Long-Term (Two-Year) Impact of <i>READ 180</i> at Grades 7 and 8 in Year 2.....	42
Differential Attrition in Immediate and Long-Term <i>READ 180</i> Impact Analyses	46
Differential Impacts of <i>READ 180</i> in MCLA and Non-MCLA Schools	47
Conclusions	52
Further Analyses	52
V. Evaluation of the Implementation of the Whole-School Intervention: Years 1 and 2.....	53
Contextual Factors in Control and Experimental MSRP schools.....	55
Professional Development Model for Teachers as Implemented	56
MCLA Course Content	57
MCLA Course Participation	59
MCLA Principal Fellowship Course Participation	63
Curriculum Resource Center (CRC) Use	63
Literacy Coaching Support	65
Coaches' <i>READ 180</i> Tasks	66
Classroom Level MCLA Implementation.....	67
Summary of Level of Implementation Attained for Whole-School Intervention	71

VI. Evaluation of the Impacts of the Whole-school Intervention: Years 1 and 2.....	73
Sample Selection	73
Data Collection	74
Data Analysis	75
Selection of Covariates	79
Treatment of Missing Data	79
Description of Samples of Teachers and Students for MCLA Impact Analyses	79
Equivalence on Teacher Characteristics.....	79
Equivalence on Baseline Teacher Indices	80
Equivalence on Student Characteristics.....	80
Equivalence on Baseline Student Achievement.....	82
Impact of MCLA on Teacher Preparedness and Frequency of Use in Years 1 and 2.....	85
Immediate Impact of MCLA in Year 1	85
Long-Term Impact of MCLA in Year 2.....	86
Differential Attrition	86
Conclusions	87
Further Analyses	87
Impact of MCLA Participation on Student Achievement in Years 1 and 2.....	87
Immediate Impact of MCLA in Year 1	87
Immediate Impact of MCLA at Each Grade in Year 1	89
Immediate Impact of MCLA at Sixth Grade in Year 2	90
Long-Term (Two-Year) Impact of MCLA at Grades 7 and 8 in Year 2.....	91
Differential Attrition in Immediate and Long-Term MCLA Impact Analyses	95
Conclusions	96
Further Analyses	96
 Summary of Year 2 RBS Data Collection Activities	 97
 References	 99

I. Executive Summary

This report presents the results from an evaluation of the first two years of the Memphis Striving Readers Project (MSRP) conducted by Research for Better Schools (RBS). MSRP, funded for five years by the United States Department of Education, comprises two interventions aimed at improving adolescent literacy and the quality of literacy instruction across the curriculum. The first MSRP intervention component, Scholastic's *READ 180* program, was implemented for two years in eight middle schools in the Memphis City Schools (MCS) district. Struggling readers in grades six through eight who met eligibility requirements were randomly selected either to participate in the supplemental *READ 180* program or serve as control group students. There were 1,740 students in the treatment and control groups at the eight MSRP schools in Year 1 (2006–2007) and 1,513 students in treatment and control groups in Year 2 (2007–2008).

The second intervention component, the Memphis Content Literacy Academy (MCLA), is a whole-school professional development program that was implemented for two years with teachers in four of the eight MSRP schools. The program, developed by University of Memphis and MCS staff, provided courses for teachers and principals, literacy coaching assistance, and instructional materials. The teacher course was designed to help teachers integrate literacy strategies into the content areas of English/language arts, social studies, mathematics, and science. In Years 1 and 2, the number of students attending MCLA schools was 2,877 and 2,408, respectively. In Years 3 and 4 of MSRP, the whole-school intervention will be implemented in the four schools that originally served as controls for this component of the study.

Researchers collected information about the implementation and impact of the two interventions using a variety of methods, including surveying, observing, and interviewing teachers as well as reviewing program documents and student scores on the Tennessee Comprehensive Assessment Program (TCAP) and Iowa Tests of Basic Skills (ITBS). The Year 1 evaluation found considerable variability in the implementation of both interventions among the school sites, partly due to delays in the delivery of equipment needed to operate the *READ 180* software programs and low participation rates in MCLA among eligible teachers. In Year 2, the *READ 180* and MCLA interventions were implemented at higher levels; however, the impact analysis failed to show immediate or long-term effects of *READ 180* on student achievement. The MCLA impact analysis concluded that intervention teachers were more prepared than teachers in non-MCLA schools to use literacy strategies but did not find evidence of impact on measures of student achievement.

Implementation and Impact of the Targeted Intervention in Years 1 and 2

The implementation evaluation of the *READ 180* intervention examined the level and variability of teacher professional development and the extent to which the *READ 180* model was implemented as planned at the classroom level as a supplement to regular English/language arts classes that all students received. To inform the implementation study, researchers analyzed information about professional development participation, data from the Scholastic Achievement Manager (SAM, which tracks student performance), surveys, and classroom observations. To

determine the impact of *READ 180* participation on student achievement, researchers analyzed TCAP and ITBS results for the treatment and control groups. The impact study explored immediate and long-term (two-year) effects of *READ 180* on student achievement.

In Year 1, RBS found wide variation in *READ 180* implementation across the eight MSRP schools. Of the 19 teachers, 13 received adequate professional development ratings and 12 were rated as adequate based on classroom observations. Overall, eight teachers/classrooms were rated as adequate and one was very close to adequate, but the remaining ten were substantially below that standard. Two factors likely affected the Year 1 implementation ratings: (1) some teachers had not received all the resources and materials necessary to fully implement *READ 180* until midway through the school year, and (2) two rounds of observations limited the strength of the results.

Drawing from SAM data, professional development attendance information, and far more (six rounds) of classroom observations **in Year 2, researchers found that the level of *READ 180* implementation in the MSRP schools improved from Year 1 to 2.** Although teacher turnover precluded calculating improvement at the classroom level, RBS found that the number of classrooms rated as adequate according to classroom observations increased from 12 (63.1%) in Year 1 to 15 (78.9%) of 19 teachers in Year 2. Although there was little change in the number of teachers rated adequate in professional development (13 in Year 1 compared with 14 in Year 2), the number of classrooms rated as adequate using SAM data increased from eight to 10 over the two years.

Analysis of the immediate and long-term impacts of *READ 180* on student achievement in Years 1 and 2 is complex. During Year 1, students eligible for the study attended sixth, seventh, and eighth grades; however, in Year 2, students previously in eighth grade had matriculated to high school while a new cohort of sixth graders enrolled in the MSRP schools. First, researchers analyzed student achievement in reading and four TCAP core subject areas in Year 1 to determine any immediate effects of first-year participation in *READ 180*. In Year 2, researchers analyzed results for 693 new, eligible sixth graders to determine if the immediate impact of *READ 180* on sixth grade students varied between years.

Researchers also tested the achievement differences between students in the control and treatment group using multi-level regression analysis that took into account student-, school-level variables such as students' socio-demographic characteristics or the percentage of low-income students attending the student's school. Researchers found no differential attrition in the treatment and control group in Years 1 and 2.

Results from the impact analyses showed that ***READ 180* participation did not have a significant effect on student achievement in reading or the four core subject areas by the end of the first year of the intervention. Year 2 findings continued to show no significant impact.** Analyzing the achievement of students who remained in the *READ 180* treatment and control groups for Years 1 and 2, researchers found significantly greater growth in mean ITBS reading test scores for seventh and eighth grade control students in Year 1, and seventh and eighth grade control students showed significantly greater growth in mean social studies scores in Year 2. No striking patterns emerged overall regarding the differences in *READ 180* impact

for the treatment or control group students. There were no immediate impacts by the end of Year 1, and only one of 21 long-term impacts was significantly different for seventh- and eighth-grade students in treatment and control groups based on a cross-sectional comparison of Spring 2008 achievement scores. Additional analyses that explore the relationship between the amount of *READ 180* instruction and its effects may shed further light on the program's impact.

Implementation and Impact of the Whole-School Intervention in Years 1 and 2

The implementation evaluation of the MCLA program examined the type and amount of professional development provided to participants, the content of the MCLA training model, the proportion of teachers who participated in different levels of professional development, and the type of literacy coaching assistance provided. Researchers also examined contextual factors affecting MCLA implementation and the professional development experiences of teachers in four control schools that did not participate in the MCLA intervention in Years 1 and 2.

The impact analyses examined teacher and student outcomes. First, researchers compared teachers' survey responses about how prepared they were to use 24 specific literacy strategies and the frequency with which they used the strategies in the classroom. Second, the analyses of the immediate and long-term impact of MCLA compared the TCAP/ITBS achievement of students at the MSRP schools. Only students who attended the schools for a majority of the instructional days were included in the analyses of immediate impact, and only those doing so for both Years 1 and 2 were included in the analysis of long-term impact.

The implementation study of the whole-school intervention used information from professional development attendance records, resource center check-out logs, literacy coaches' daily logs, surveys, interviews and materials such as syllabi and course templates that guided instructors who taught the professional development classes. Researchers assigned each of the four MCLA schools an implementation rating that was the arithmetic average of the following scores: teachers' levels of course attendance, principals' MCLA involvement, the number of resources teachers borrowed from the MCLA resource library, and the amount of coaching assistance teachers received.

Results show that the level of MCLA implementation at the four schools increased from Years 1 to 2. Whereas attendance was uneven in the Year 1 principal fellowship course, all four principals attended all classes and relevant MCLA events in Year 2. Teacher attendance in the professional development course was high: 65.6 percent of teachers who completed both semesters of Year 2 classes had attended 80 percent or more of the classes. It appeared that fewer teachers used the curriculum resource center (CRC) provided by the MSRP grant in Year 2; however, comparisons with Year 1 results are limited since only three of four MCLA schools provided information on CRC usage in Year 2.

The analysis of the immediate effect of MCLA on teachers found that teachers in MCLA schools were significantly more likely than teachers in control schools to feel prepared to use certain literacy strategies ($p < 0.05$). The estimated long-term impact of MCLA on teachers' preparedness to use targeted literacy strategies was sustained in Year 2. While there was a significant difference in favor of the teachers in MCLA schools with regard to the frequency of strategy

implementation in Year 1, this difference was not sustained in Year 2. **There was no significant difference in the reported frequency of implementing various strategies between teachers in control and MCLA schools at the end of Year 2.**

In Year 1, impact analyses detected no immediate effects of MCLA on students' achievement as measured by the TCAP and ITBS. **Results in Year 2 showed no significant impact on students in MCLA schools**, and a longitudinal analysis of the two-year effect on achievement produced a mixed bag of results, with some findings favoring the control group students. Future analyses that explore MCLA impact among students whose teachers participated at high intervention levels or among students who were enrolled in multiple classes taught by trained teachers may provide deeper insight into the effects of MCLA.

II. Introduction and Study Background

Introduction

This report presents the results from an evaluation of the first two years of the Memphis Striving Readers Project (MSRP) conducted by Research for Better Schools (RBS). MSRP, funded for five years by the United States Department of Education, comprises two interventions aimed at improving adolescent literacy and the quality of literacy instruction across the curriculum. Eight MCS middle schools were chosen for inclusion in the study. Table 1 provides the names and enrollments of these eight schools.

Table 1: Student Enrollments in Schools Participating in Striving Readers Study

Memphis Middle Schools Participating in MSRP	Enrollment in 2006–2007 ^a	Enrollment in 2007–2008 ^a
A. Maceo Walker	856	724
American Way	997	1034
Corry	413	374
Hamilton	635	520
Hickory Ridge	858	856
Lanier	640	603
Riverview	471	405
Sherwood	915	759
Total	5785	5275

^a Average Daily Membership reported in school report cards on the Tennessee Department of Education website.

The first intervention component, Scholastic’s *READ 180* program, has been implemented for two years in the MSRP schools. Students in the sixth through eighth grades who demonstrated the strongest need for reading support (i.e., performed in the bottom reading quartile of a prior Tennessee Comprehensive Assessment Program [TCAP]) were randomly selected to participate in the supplemental program.¹ The *READ 180* intervention served 698² students in the eight schools in Year 1 (2006–2007) and 289 new sixth-grade students were assigned to the intervention in Year 2 (2007–2008).

The second intervention component, the Memphis Content Literacy Academy (MCLA), is a whole-school professional development program that was implemented for two years with teachers in four of the eight MSRP schools. The program, developed by University of Memphis and MCS staff, was designed for teachers in the English/language arts (ELA), social studies, mathematics, or science content areas or special education resource teachers (who do not teach in self-contained classrooms). In Years 1 and 2, the number of students attending MCLA schools was 2,877 and 2,408, respectively. In Years 3 and 4 of MSRP, the whole-school intervention will be implemented in the four schools that originally served as controls during the first two years of this study.

¹ As noted on page 29, students receiving special education services were omitted from all impact analyses.

² Data sources: For the targeted intervention enrollment, the “Enrollment file 06-14-07 SR Schools Only” data file provided by MCS. For the school populations, data were downloaded on July 8, 2008, from <http://www.memphis-schools.k12.tn.us/admin/communications/directoryofschools.html>

Researchers collected information from the targeted and whole-school interventions through surveys, observations, document reviews, computer-generated data, interviews, and focus group sessions. The Year 1 evaluation found considerable variability in the implementation of both interventions due to delays in the delivery of equipment needed to operate the *READ 180* software programs and low participation among teachers in the professional development program at two of the four intervention schools. By Year 2, the *READ 180* and MCLA interventions were implemented at higher levels; however, an impact analysis presented in this report failed to show immediate and long-term effects of *READ 180* on student achievement. The MCLA impact analysis, also summarized in this report, concluded that intervention teachers were more prepared than teachers in non-MCLA schools to use literacy strategies.

Background Context

The United States Department of Education awarded MCS a five-year Striving Readers grant to help address the city's significant educational needs. MCS, which serves more than 110,000 students, is the 21st largest K–12 district in the United States. Over 95 percent of the 196 MCS schools are Title I schools, and 71 percent of students qualify for free or reduced-price lunches (The Urban Child Institute, 2008). Approximately 86 percent of MCS students are African American, 8 percent are white, and 6 percent are other races and/or ethnicities (MCS, 2008). Other data show that 71 percent of students in the sixth through eighth grades scored below the 50th percentile on the Reading/Language Arts portion of TCAP (Potts, Perkins, Heeren, Harris, and Feldman, 2008) and that only 69.2 percent of MCS students graduate high school within four years (Hart, 2008).

Theoretical Rationale for and Description of the Intervention Models

Description of the Targeted Intervention

READ 180 is a commercially available reading intervention program from Scholastic that targets struggling readers in the fourth through twelfth grades. The Enterprise Edition is the most recent version, and it combines a software program, teacher-directed instruction using a textbook and similar resources, and independent or modeled reading (i.e., reading while listening to audiobooks). Close adherence to the structure of the program requires 90 minutes divided into four 20-minute and one 10-minute blocks. The first 20 minutes and a final 10-minute wrap-up involve whole-group instruction. The other three 20-minute blocks require students to rotate between teacher-led small-group instruction, individual use of the proprietary *READ 180* software, and reading leveled fiction and nonfiction texts provided with the program.

Students Targeted by the Intervention

MCS created a pool of struggling readers by identifying students at all eight MSRP schools who scored in the lowest quartile on the Reading and Language Arts section of the TCAP exam. In fall of 2006, all identified students were randomly assigned to the control or treatment condition, and in fall 2007, treatment students who were still enrolled in MSRP schools (that is, students who moved from sixth to seventh grade or seventh to eighth grade) were again enrolled in

READ 180 Also, during the fall of 2007, incoming sixth-grade students who were identified as struggling readers were randomly assigned to the control or experimental condition. During the 2007–2008 school year, 480 students in the sixth, seventh, and eighth grades were enrolled in *READ 180* for the majority of the school year, and 942 students comprised the control group.³

Students assigned to *READ 180* experience the intervention in addition to their regular language arts classes and other courses related to language arts (e.g., reading, creative writing) that are offered at their schools. Control students experience the same language arts classes, reading classes, and/or other classes related to language arts that would be offered in their schools if MSRP did not exist.

Logic Model for Targeted Intervention

The logic model for the targeted intervention, as published in Scholastic’s *READ 180 Enterprise Edition Research Protocol and Tools* (2007), appears as Figure 1. The *READ 180* “package” purchased for the 2007–2008 school year by MCS included the following:

- networking applications to link the student software with reporting software that teachers and administrators can use
- videos related to the readings in the textbook intended to build knowledge and spark discussion
- teacher materials including additional lessons and strategies and instructions for using the materials and reporting software
- the classroom software, leveled texts, and *rBooks* or *flexBooks* (or workbook-type textbooks) for student use
- assistance and training from Scholastic personnel for teachers, administrators, and instructional technology professionals
- formative observations of classroom teaching, by Scholastic personnel, with suggestions for improvement

Graphics illustrating the instructional model and detailing the rotation activities appear as Figures 2a and 2b.

Professional Development Model Components

The professional development component includes the only changes in implementation between Years 1 and 2. In both years, new *READ 180* teachers were expected to attend two all-day training sessions (experienced teachers were expected to attend at least one). Additionally, the district hosted “networking meetings” (four during Year 1 and seven during Year 2) in which Scholastic representatives taught teachers how to use or improve different components of *READ 180*, such as using the data generated by the *READ 180* software to differentiate instruction or using supplemental *READ 180* teaching materials for strategic vocabulary instruction.

³ There was some attrition and a few instances in which students were opted out of the intervention; these are detailed in the “Sample Selection” section of part IV. These enrollment numbers differ from those in the section describing the impact of *READ 180* because the impact numbers include all students who were in the original design, and the numbers in this section include only those students who were actually enrolled for most of the school year.

MCS encouraged teachers to complete Scholastic’s “Best Practices for Reading Intervention,” a seven-part online course designed for *READ 180* classroom teachers (during Year 2, MCS provided a small monetary stipend to teachers who did complete the online course). Finally, in Year 2, Scholastic representatives conducted one classroom observation per teacher and provided feedback to help improve the teaching of *READ 180*. (Scholastic provides detailed descriptions of the all-day implementation training sessions and the online courses at, respectively, http://teacher.scholastic.com/products/read180/prof/implement_train.htm and <http://teacher.scholastic.com/products/read180/prof/bestpractices.htm>.)

Classroom Model Components

The planned instruction model was to follow the published and recommended *READ 180* model. According to the *Leadership Implementation Guide: Supporting READ 180 in Your District* (2005) published by Scholastic, the recommended class size for *READ 180* is 21 or fewer students. Scholastic recommends that students be divided into three homogenous groups according to diagnostic assessments and regrouped as assessments indicate. (According to Scholastic’s *READ 180* training materials, Scholastic allows for alternate grouping strategies, such as purposefully creating heterogeneous groups or considering behavioral issues to guide grouping.) Students are to be in *READ 180* class for 90 minutes during every school day. Figures 2a and 2b detail Scholastic’s recommendation for use of instructional time and provide some details about the targeted areas of reading and instructional approaches.

READ 180 software provides instruction in decoding and word recognition, spelling, fluency, vocabulary, and comprehension. Whole-group and small-group instruction include a variety of instructional approaches, including fluency exercises, question stems, use of graphic organizers, activation of prior knowledge, and cooperative group work (among others). The Scholastic Achievement Manager (SAM) automatically generates student-level data based on work students have done and assessments they have completed using the *READ 180* software.

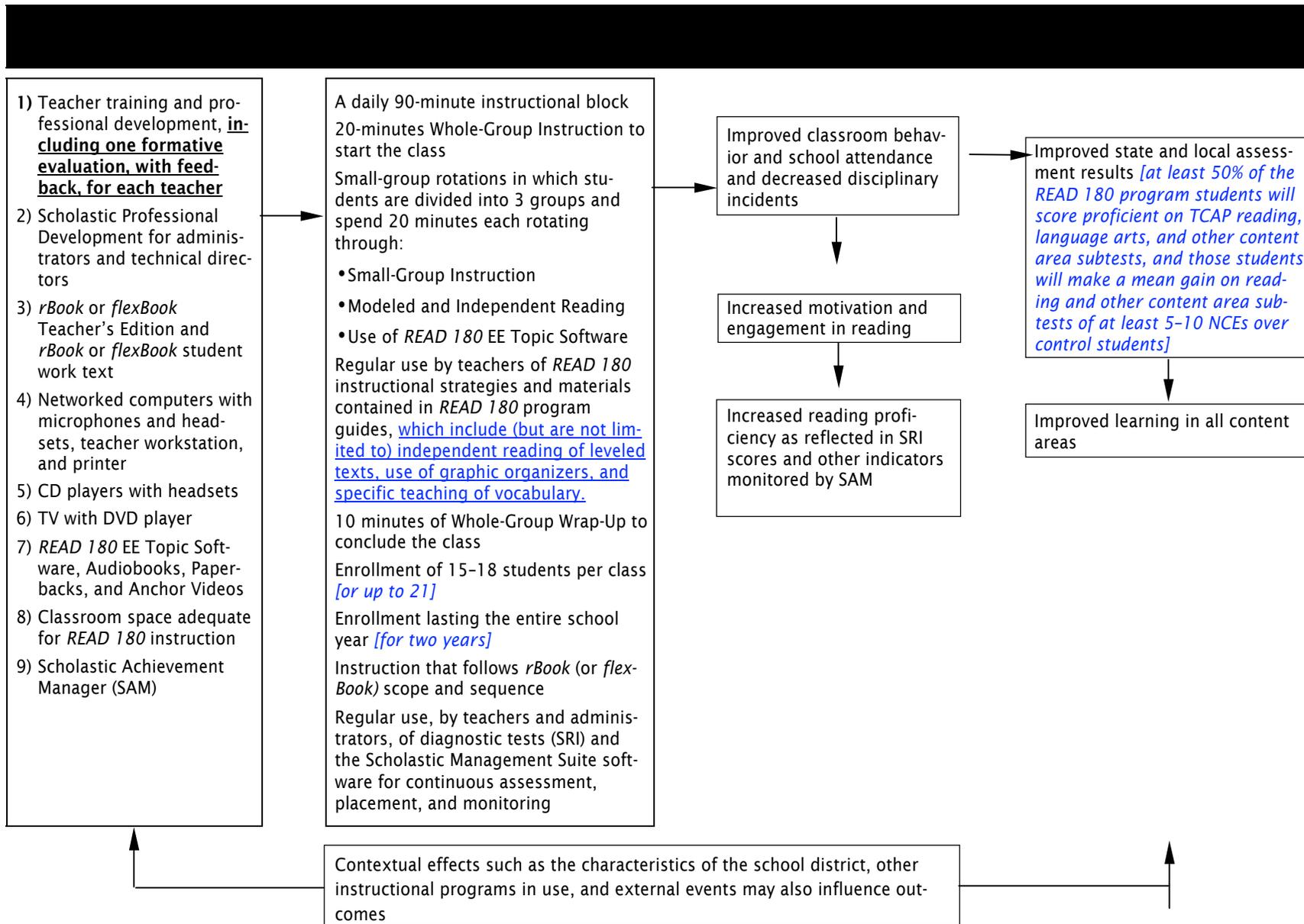
All students in *READ 180* classes are provided with a *flexBook*,⁴ the course textbook. All *READ 180* classrooms have libraries with a variety of fiction and nonfiction leveled texts provided for the modeled and independent reading rotation of *READ 180*. During the *READ 180* software rotation, each student has individual access to a computer to complete the exercises. Teachers use their dedicated computers for recordkeeping and for tracking student progress through the instructional software.

READ 180 includes a number of assessment tools in its software. Regular reports of student progress through the instructional software are available to teachers through the networked computers using SAM and the Scholastic Management Suite (SMS) software. Teachers are expected to administer the Scholastic Reading Inventory (SRI) three times per school year. Additionally, the teachers are expected to use the data from SAM and SMS, the SRIs, and other assessments that might be chosen by the teacher to determine whether lessons are working, to differentiate instruction, and to regroup the students.

⁴ During the 2006–2007 school year, students used the *rBook*. The *rBook* and the *flexBook* are the two versions of the same text, with similar lessons, exercises, strategies, etc. The versions have different reading selections so students who are in *READ 180* for two years do not experience identical readings over those two years.

Figure 1: Logic Model of Targeted Intervention

The READ 180 Enterprise Edition Logic Model



Logic Model copyright © 2007 Scholastic Inc. *Text in blue italics is specific to MSRP. Blue underlined text was added for clarity by RBS. Bold underlined text describes the only differences in the targeted intervention between project Years 1 and 2.*

Figure 2a: READ 180 Instructional Model



Figure 2b: Description of READ 180 Rotation Activities

Small-Group Rotations				
Whole-Group Direct Instruction	Small-Group Direct Instruction	<i>READ 180</i> Software	Modeled and Independent Reading	Whole-Group Wrap-Up
Using the <i>READ 180</i> instructional materials, the teacher begins the day by providing systematic instruction in reading, writing, and vocabulary to the whole class.	Using the <i>rBook</i> and <i>Resources for Differentiated Instruction</i> , the teacher works closely with students so that individual needs can be met.	Students use the software independently, providing them with intensive, individualized skills practice.	Students build reading comprehension skills through modeled and independent reading of the <i>READ 180</i> paperbacks and audiobooks.	The session ends with 10 more minutes of whole-group instruction.

The above graphic and table were copied on November 8, 2007, from <http://teacher.scholastic.com/products/read180/overview/instrmodel.htm#small-group>

Theoretical Rationale for and Description of Whole-School Intervention Model

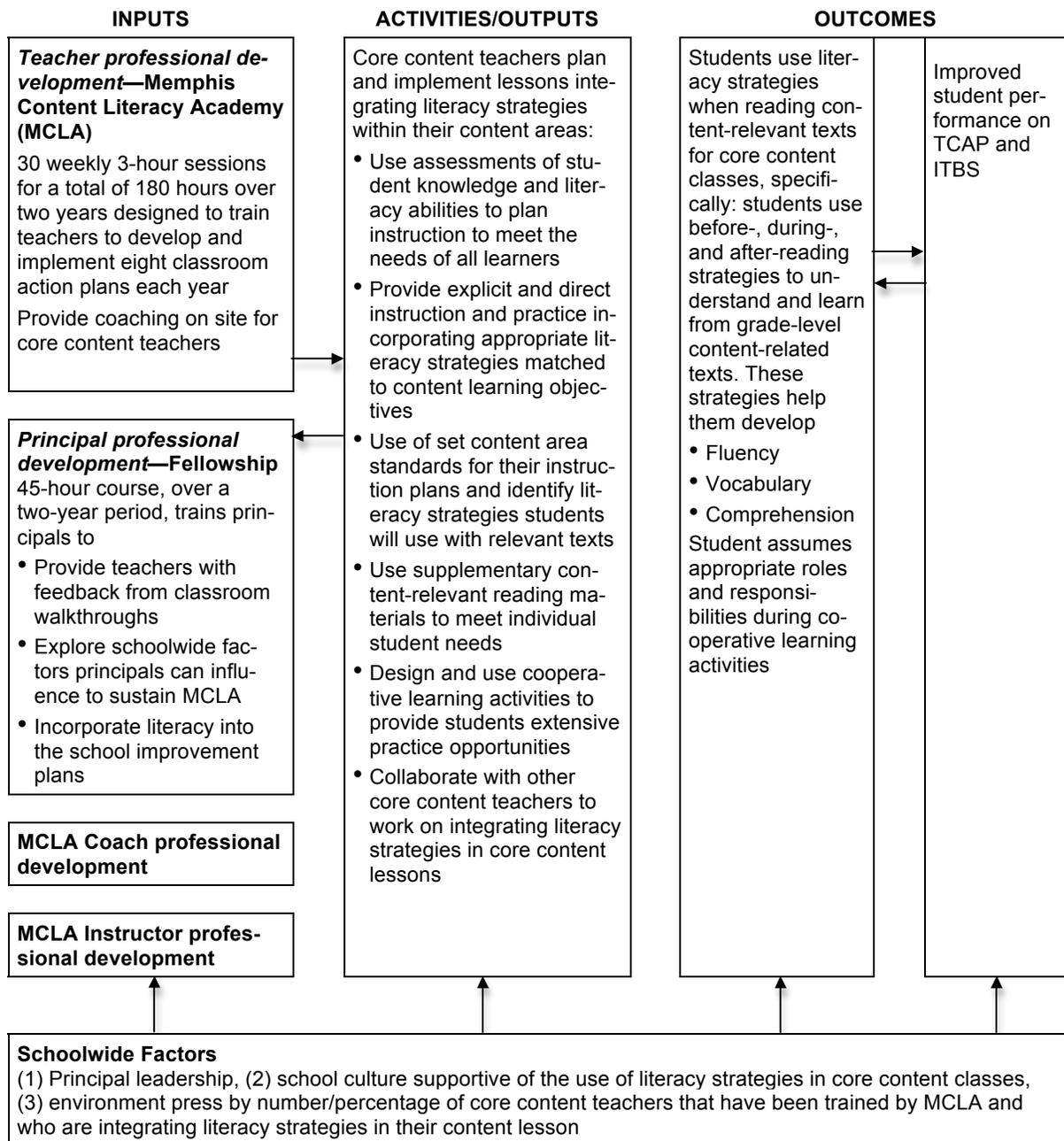
MCLA is a whole-school intervention designed to improve teaching and learning through intensive professional development, onsite literacy coaching assistance, a leadership seminar, and use of grant-funded curricular resources. Four of the eight participating schools were randomly assigned to receive the whole-school intervention during Years 1 and 2 of the study. During Years 1 and 2, MCLA was implemented at A. Maceo Walker, Hamilton, Riverview and Sherwood middle schools. In Year 1, 145 full-time content area teachers were invited to attend two years of classes that would focus on infusing literacy into ELA, mathematics, science, and social studies lessons. Special education resource teachers (those not teaching self-contained classes) were also invited to join MCLA in Year 2. Teachers from the other four schools will be able to participate in MCLA during Years 3 and 4.

Developers hypothesized that greater and more effective integration of literacy strategies by teachers would lead to student performance improvements in reading and core academic content classes. The intervention was designed so that teachers with no prior knowledge of or experience with literacy integration had opportunities to practice strategies in the MCLA course and then, with coaching support, gradually assumed responsibility for helping students internalize those techniques. Two tools designed for the evaluation elaborate on this theory of action: the first is a logic model of intervention activities, outputs, and anticipated outcomes; the second is a significantly more comprehensive and in-depth rendering of the intervention, an “Innovation Configuration Map,” created in partnership among evaluators, program developers, and project implementation leaders and staff.

As Figure 3 shows, developers planned to offer content-related materials on literacy integration in the evening course and in a \$40,000 curriculum resource center (CRC) stocked with content-rich, multi-leveled materials and an onsite literacy coach to assist with strategy implementation. As a result of participation, developers anticipated that teachers would gain a deeper understanding of the need for literacy integration into the content areas and use a series of research-based literacy integration skills with increased frequency and confidence.

Developers also hoped that the intervention would create a sharpened schoolwide focus on literacy as teachers shared techniques and experiences and assessed the program’s impact on their students. Ultimately, developers anticipated that the transformed literacy-saturated middle school environment would boost student achievement in reading as well as in core content-area classes.

Figure 3: Logic Model of the Memphis Striving Readers Whole-School Intervention



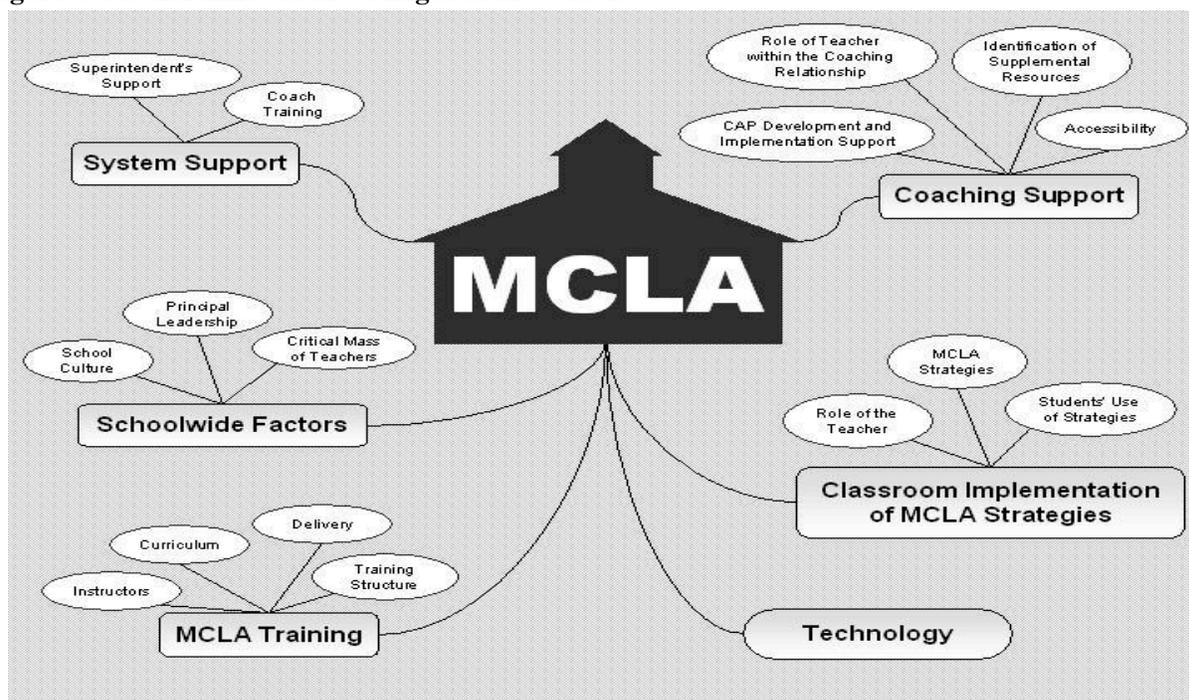
The Innovation Configuration Map

Evaluators met with university and school district partners in Years 1 and 2 to design an Innovation Configuration (IC) Map (Hall and Hord, 2006) that explicated the intervention’s main constructs. Dr. Gene Hall, co-creator of the IC Map concept, co-facilitated the meetings. The map describes the “ideal” implementation of each intervention component and the variations that describe a range of levels of implementation fidelity. Ideally, the IC map will be used to reflect on teacher practice, monitor classroom implementation, identify the areas of a program in most

need of attention, and develop evaluation instruments. See Appendix N-1 for a full version of the IC Map.

Figure 4 presents a pictorial overview of the six major domains detailed in the IC Map as conceptualized after the first of two years of implementation. In addition to providing professional development, MCLA developers had planned to elicit support from the superintendent, use technology to share implementation strategies (a component that was later dropped from the model), and train the literacy coaches so that they were able to mentor teachers and deliver high-quality assistance. Developers attempted to recruit a critical mass of participants, defined, ideally, as all core content teachers in each school.

Figure 4: MCLA Innovation Configuration Overview



Professional Development Model Components

The MCLA teacher and principal course syllabi, coach job descriptions, and an inventory of instructional materials best describe the intervention's four main components. The teacher course was designed to meet weekly for three hours over two years, for a total of 180 hours of professional development. According to the original proposal, planners had hoped to retain 80 percent of approximately 120 eligible content-area teachers who would select one of four evening content courses (mathematics, science, ELA, and social studies) according to their primary teaching assignment. Developers offered a similar course structure across content areas that exposed participants to discipline-specific materials and wrote templates for the instructors (highly qualified reading and content-area specialists) to follow when teaching the course to ensure standardization. The developers' goal was for participants to see strategies modeled in the course and then, with a coach's guidance, practice using those strategies in their own classrooms. By

issuing approximately ten “classroom action plan” (CAP) course assignments over two semesters, designers hoped that teachers would be compelled to work with coaches in school on a weekly basis. Finally, the grant team also proposed supplementing MCLA training by recording and sharing videotapes of participants to allow teachers to observe common issues encountered in implementing the literacy strategies in MCS content classrooms.

Developers aimed to bolster schoolwide support for MCLA by also providing a principal fellowship course to allow school principals to brainstorm possible solutions to commonly experienced problems, apprise them of the details of MCLA, and guide them in mentoring teachers who were implementing MCLA strategies. Initially, the fellowship was designed to be held weekly for three hours during the fall semester; however, since Year 1 weekly attendance targets were unmet, developers cast a wider net in Year 2 and invited assistant principals and other school administrators, as well as principals, to attend less frequent sessions (once per month) across both semesters of the 2007–2008 school year.

Assistance provided by onsite literacy coaches comprises the third component of MCLA. In Years 1 and 2 of the program, the literacy coaches—individuals with five years of teaching experience, a Master’s degree, and a strong literacy background—were responsible for observing, monitoring, and assisting teachers with meeting the weekly objectives of the CAPs; monitoring *READ 180* program delivery; visiting regularly with principals to keep them informed of teachers’ needs; and maintaining the CRC. In Years 1 and 2, coaches were also expected to participate in a wide range of professional development activities provided by MCLA developers in the areas of mentorship, urban education, adolescent literacy, and in Scholastic’s *READ 180* program.

The CRC comprises the fourth and final component of the MCLA professional development program. Housed in each of the participating schools, each CRC contains leveled books, kits, reference books, and other materials to assist teachers with integrating literacy into their content area classes. Evening course instructors and onsite literacy coaches promoted the use of CRC materials, which remained in the schools when MCLA ended.

Classroom Instruction Model Components

MCLA designers identified twelve key instructional strategies to be the primary foci of the teacher course prior to implementation, including the use of graphic organizers, comprehension monitoring techniques, question generation, repeated oral reading, preteaching vocabulary, and direct, explicit instruction. Course syllabi in Years 1 and 2 required teachers to use strategies aimed at improving students’ vocabulary, fluency, and comprehension and to meet with coaches to plan lessons designed to bolster strategy implementation. The MCLA course instructional model was designed to provide teachers opportunities to practice modeling the literacy strategies among colleagues as part of training before implementing them with their students. Participants were then expected to model the strategies at school and help students to adopt and use the strategies independently and appropriately. Developers had hoped that teachers would integrate the strategies into existing class activities rather than view them as separate and distinctive lessons. Although developers did not plan a formal system for using student data to inform deci-

sion-making in MCLA, they built into the CAP assignments an informal student assessment procedure.

Changes Made to the Professional Development Model between Years 1 and 2

During the first two years of MCLA, developers fine-tuned service delivery based upon experience gained from implementation and the feedback of participants and evaluators. The following changes, described in more detail in subsequent sections of this report, were made to MCLA between Years 1 and 2:

- Inclusion of special education resource teachers in the teacher course
- Expansion of principal fellowship to include other school administrator leaders and new seminar schedule
- Replacement of two literacy coaches who resigned or were promoted to other positions
- Modification of the professional development course format to rely more heavily on small group activities and class presentations and less on lecture
- Incorporation of “before, during, and after” reading strategies in the CAPs

Based on lessons learned from Year 1, teacher classes were organized into three-week cycles: the first week focused on an explication of key strategies, the second week allowed teachers to model the strategies through presentations, and the third week offered them opportunities to work collaboratively with colleagues to develop their classroom action plans (CAPs). This format represented a slight departure from Year 1, during which MCLA instructors covered a greater number of literacy strategies. During the final semester of Year 2, participants reviewed all strategies covered during the first three semesters.

Brief Overview of Key Evaluation Design Features

The evaluation design for the MSRP study addresses the impact of *READ 180* and MCLA on student and (in the case of MCLA) teacher outcomes. The measures of student outcomes are the Iowa Tests of Basic Skills (ITBS) to measure reading achievement, and the TCAP, which measures achievement in reading/language arts, mathematics, science, and social studies. Outcome measures from the ITBS include the total reading standard score and the vocabulary and comprehension subtest standard scores. Outcome measures from the TCAP include scale scores in the four content areas.

There are two teacher outcome measures: an index of the teacher’s perceived preparation to employ literacy strategies in the classroom and an index of the teacher’s perception of how frequently these strategies are employed. Each index is based on ratings of preparation or frequency for 24 literacy strategies.

***READ 180* Impact**

The primary purpose of the evaluation of the targeted intervention is to determine the impact of *READ 180* on struggling readers’ reading achievement, as measured by the ITBS, and on their achievement in reading/language arts, mathematics, science, and social studies, as measured by TCAP. Since the targeted intervention will be implemented for four years, its impact on student

achievement will be evaluated after one year (immediate impact) and after one or two additional years (long-term impact). Annual variations in both immediate and long-term impacts are also of interest. Finally, variations in immediate and long-term *READ 180* impact between MCLA and control schools during Years 1 and 2 will be examined.

The evaluation design for the targeted intervention is intended to answer the following research questions:

- What is the immediate, one-year impact of *READ 180* on the reading and content area achievement of struggling readers at the end of the first year of participation?
- Does the immediate impact of *READ 180* on the reading and content area achievement of struggling readers in the sixth grade vary over the four study years?
- What is the long-term, two-year impact of *READ 180* on the reading and content area achievement of struggling readers?
- Does the long-term impact of *READ 180* vary over the four study years?
- Do the immediate and long-term impacts of *READ 180* vary between schools receiving and not receiving the whole-school intervention, MCLA, during the first two years of the study?

The impact evaluation of the targeted intervention is based on the randomized assignment of students. In order to evaluate the immediate impact of *READ 180* in Year 1, all eligible struggling readers in the sixth to eighth grades were randomly assigned to control and treatment groups within each of the eight schools. The treatment students participated in *READ 180* during Year 1; the control students did not.

The immediate impact of *READ 180* in Year 1 was determined by comparing the achievement levels of these two groups at the end of the year. The immediate impact of *READ 180* in Year 2 was determined by randomly assigning eligible sixth-grade struggling readers in Year 2 to treatment and control groups for the targeted intervention within each school. The immediate impact on these sixth-grade students in Year 2 was compared to the impact on sixth-grade students in Year 1. In order to evaluate the long-term impact of *READ 180* in Years 1 and 2, the Year 1 treatment students who remained enrolled in one of the eight participating schools in Year 2 continued to participate in *READ 180*. The control students continued not to participate. The long-term impact of *READ 180* was determined by comparing the achievement levels of sixth- and seventh-grade treatment and control students who continued to be enrolled in one of the eight participating schools in Year 2 in the seventh and eighth grades.

MCLA Impact

The primary purpose of the evaluation of the MCLA whole-school intervention is to determine the immediate and long-term impact of MCLA on struggling readers' reading achievement, as measured by the ITBS, and on their achievement in reading/language arts, mathematics, science, and social studies, as measured by the TCAP. Since MCLA is a whole-school intervention, this evaluation is also designed to determine the impact of MCLA on all students in participating middle schools. Another purpose of the evaluation of the whole-school intervention is to determine the immediate and long-term impact of MCLA on teachers' perceptions of

their use of literacy strategies as measured by a teacher survey designed by Feldman and Feighan (2007) with input from MCLA developers.

The evaluation design for MCLA is intended to answer the following research questions:

- What is the immediate, one-year impact of MCLA on teacher preparation for and use of instructional strategies that promote student literacy?
- What is the long-term, two-year impact of MCLA on teacher preparation for and use of instructional strategies that promote student literacy?
- What is the immediate, one-year impact of MCLA on the reading and content area achievement of struggling readers and of all students enrolled in MSRP schools?
- Does the immediate impact of MCLA on the reading and content area achievement of struggling readers and of all students in the sixth grade vary between the first and second year of its implementation?
- What is the long-term, two-year impact of MCLA on the reading and content area achievement of struggling readers and of all students enrolled in MSRP schools?

The impact evaluation of the whole-school intervention is based on the randomized assignment of schools. The eight schools participating in the MSRP study were grouped into four matched pairs, and, within each pair, one school was randomly assigned to receive the whole-school intervention, MCLA, during Years 1 and 2, and the other was assigned to be a control school. In Years 3 and 4, the four control schools will receive MCLA.

The evaluation of the impact of MCLA on students was based on students in the sixth through eighth grades who were enrolled in a treatment or control school for a majority of the instructional days. Details of how these students were identified and the number who met this condition are described in Part VI of this report.

The immediate impact of MCLA was determined by comparing the achievement levels of students in the treatment and control schools in Year 1. The immediate impact of MCLA on sixth-grade students enrolling in treatment and control schools in Year 2 was also determined and compared to the impact on sixth-grade students in Year 1. The long-term impact of MCLA in Years 1 and 2 was determined by comparing the achievement levels of sixth- and seventh-grade students in treatment and control schools during Year 1 who continued to be enrolled in one of the schools from the same design group (treatment or control) in Year 2 in seventh or eighth grade.

III. Evaluation of the Implementation of the Targeted Intervention, Years 1 and 2

Summary of the Design

Because this originally was considered an effectiveness study rather than an efficacy study, a plan to develop detailed implementation ratings of the targeted intervention was not created for Year 1. The plan to study implementation more closely and develop classroom implementation

ratings for Year 2 was generated during Year 2, as feedback about the appropriate role of researchers in collecting and reporting these data from the technical assistance provider became available. The research questions that were developed are as follows:

1. What were the levels and variability of implementation of teacher professional development in Years 1 and 2?
2. What were the levels and variability of implementation at the classroom level in Years 1 and 2?

Table 3 lists the research questions and indicates the relevant data that are available from Years 1 and 2. As in Year 1, data from attendance records and developer materials were used to inform professional development implementation, and Scholastic Achievement Manager (SAM) and observation data were used in the calculation of classroom implementation ratings. However, in Year 2, classroom observations were conducted six times (improving on the two times in Year 1): thrice by the evaluator, twice by MCS, and once by Scholastic. Also, student surveys were administered by MCS during Year 2, and those data were made available to the evaluator. Unfortunately, the response rate for the teacher surveys was only about 50 percent, so these data were not used to calculate implementation ratings, although they may be used for background or context for secondary or exploratory analyses. All data continue to be available at the classroom level, but not at the individual class period level.

Development of the Ratings and Scale for Years 1 and 2

As noted in Table 2, the sources of data for rating the implementation fidelity of *READ 180* included teacher and student surveys, classroom observations, data generated by SAM, and documents related to professional development (copies of the survey and observation instruments can be found in Appendix M). Findings from all of these sources were translated to a 4-point scale ranging from 1 to 4.⁵ For all ratings, “adequate” is defined at 3 or above—the “moderate” or “high” level. The “Professional Development Scales” and “Levels and variability of implementation at the classroom level” sections each include more detailed descriptions of the specific data sources used for those areas. Table 3 provides the scale and indicates in gray those that are considered “adequate.”

Table 2: Rating Scale for Teacher Professional Development and Classroom Implementation

Scale Score	Description
4	High
3	Moderate
2	Low
1	Minimal

It should be noted that these are correctly identified as classroom ratings, rather than teacher ratings. Teachers do not always control whether the class they are teaching is on model. For ex-

⁵ The Year 1 Executive Summary of Implementation submitted in August 2008 included presented classroom ratings on a four-point scale that ranged from 0 to 3. These numbers have been changed to match the scales of Year 2 and whole-school implementation ratings.

ample, if the *READ 180* computer server is down or the students are taking the TCAP, the average number of sessions per week that students log on to the computers will drop. Similarly, several professional development opportunities are frequently scheduled on the same days just before the school year begins, and at least one teacher was required by his or her administrator to attend, instead of the all-day *READ 180* training, an all-day session related to special education services.

Table 3: Years 1 and 2 Data Sources Linked with Implementation Research Questions—Targeted Intervention

Research Questions	Measures/Data Sources							
	Surveys		SAM	Observations			Record Review	
	Teacher [†]	Student	Developer/ District	Evaluator	Developer	District staff	MCS—PD Attendance & online course completion records	Developer documentation and handouts
What were the levels and variability of implementation of teacher professional development in Years 1 and 2?								
Types/amount of professional development provided to teachers							Yrs. 1&2	Y2
Proportion of teachers at different levels of professional development	Y1						Yrs. 1&2	
Proportion of teachers at adequate level of professional development	Y1						Yrs. 1&2	
Types/amount of professional development provided to district leaders								
Proportion of leaders at different levels of professional development								
What were the levels and variability of implementation at the classroom level in Years 1 and 2?								
Proportion of classrooms supplied with materials, resources, and technology	Y1	Y2		Yrs. 1&2	Y2	Y2		
Classrooms in which model was implemented at different levels	Y1	Y2	Yrs. 1&2	Yrs. 1&2	Y2	Y2		
Classrooms in which model was implemented at adequate level or above	Y1	Y2	Yrs. 1&2	Yrs. 1&2	Y2	Y2		

Year 1 Implementation Study

Details of the Year 1 implementation study and associated classroom ratings were first reported in August 2008. They are presented here in their entirety so readers can see the differences in the amounts and types of data gathered between Years 1 and 2 and the resulting differences in the development of the ratings.

Professional Development Levels

The professional development participation score was developed by adding together the points assigned to different types of professional development. There were four types of professional development: (1) attendance at each (of three) all-day session earned a 2, (2) attendance at each (of six) networking meeting earned a 1, and (3) each year of experience teaching *READ 180* earned a 2 (up to a maximum of 6). This resulted in a possible total of 18 points. Evaluators, in consultation with MCS staff members, determined that professional development scale scores greater than or equal to ten would be considered “excellent” and be assigned a “4” on the ratings scale, those between seven and nine would be considered “good” (and rated a “3”), those between four and six would be “moderate” (“2”), and scores three and below would be “poor” (“1”). These ratings are included in the second column of Table 4.

Levels and Variability of Implementation at the Classroom Level

Implementation fidelity and variability were monitored through classroom observations, teacher surveys, and data generated by the SAM, which tracks the progress of students and the use of *READ 180* tools by teachers. As this section describes the implementation of *READ 180*, not its impact, the variables used focused on what was made available to or completed by students, not how well students performed on tasks or assessments. For example, the Scholastic Reading Counts! (SRC) variable included is the number of quizzes that students took (which is a rough estimation of how many books a student read during the independent reading portion of the *READ 180* class), not how well they did on those quizzes.

Evaluators first examined all included variables and created an equation for translating each survey, SAM, or observational variable to the 4-point scale. Second, an equation was created that encompassed the data from within each source of data (surveys, SAM, observations, and professional development). Finally, those scores were averaged to create the ratings that appear in Table 4. When data were missing, they were left out of the second (or within-source) and overall equations. For example, if the May 2007 observation did not happen, the February 2007 observation rating served as the average observation rating. If there was no survey linked with a specific teacher/classroom ID, the overall rating was calculated using only professional development, SAM, and observation data. (The equations used to calculate the ratings are included in Appendix M.)

Members of the evaluation team completed classroom observations during February and May of 2007. All 19 *READ 180* teachers were observed at least once; 11 teachers were observed during both February and May. If a teacher was observed twice, ratings from these two sets of observations were calculated separately and averaged; if a teacher was observed once, the rating is from just that observation. Observations focused on the extent to which teachers structured the

class and the lessons observed according to the *READ 180* model and the extent to which classrooms had the resources and materials required for the program.

The observation protocol used during February 2007 included a rubric that provided descriptions of different levels of implementation and asked observers to rate the environment and lessons presented. The ratings used for this report are these:

- Environment: schedule, i.e., to what extent the class followed the 90-minute model with 20 minutes of whole-group instruction, 20 minutes each of small-group instruction, computer use, and independent/guided reading, and a 10-minute wrap-up
- Environment: room arrangement, i.e., the extent to which the room and furniture are arranged appropriately for the *READ 180* program rotations
- Presentation, time, and content of whole-group instruction
- Presentation, time, and content of small-group instruction
- Use of the Scholastic *rBook* for instruction
- Use of instructional software and length of time software was used
- Engagement of students in independent reading
- Presentation, time, and content of whole-group wrap-up

The evaluation team completely re-created the observation protocol between February and May of 2007. Items on the observation protocols used during May were more specific in the information required. For example, observers were asked to record the levels of engagement of students (or the extent to which the students were on task) during the different rotations of the class. The items used for the ratings in this report include the following:

- Time (in minutes) of different portions of class
- Presentation of whole-group instruction multiplied by student engagement in whole-group instruction
- Presentation of small-group instruction multiplied by student engagement in small-group instruction
- Use of individual computers multiplied by student engagement in computer rotation
- Engagement in independent reading
- Presentation and content of whole-group wrap-up multiplied by student engagement in whole-group wrap-up
- Number of students in class (was the number 21 or fewer, as specified by the model)
- Use of Scholastic books and materials
- Room space and arrangement of furniture

Surveys were administered during the summer after Year 1; 14 teachers completed surveys (one additional teacher completed a survey but did not provide any way of identifying herself or her classroom). Survey questions focused on availability and use of specific products, equipment, and materials. (The survey also asked about the number of years a teacher had taught *READ 180* before Year 1; this was included in the professional development scale). Items related to the following topics were included in the survey rating:

- Teacher use of Red Routines (Scholastic lesson plans for *READ 180*), SAM reports, and purposeful strategies for forming and re-forming small groups
- Month that CD player and *READ 180* teacher supplies were received
- Month that computers and software were received
- Frequency of availability of working computers, software, and other technology
- Month that *rBooks* were received

SAM data were generated for all students in *READ 180*, so evaluators linked all students with their teachers and averaged together the data from all students taught by each teacher. The SAM variables used for this report included the following:

- Average number of *READ 180* computer software sessions per week
- Average daily number of minutes spent in these sessions
- As noted in Table 4, eight out of 19 *READ 180* classrooms (42.1%) were rated as “adequate,” and one additional classroom was less than one-tenth of a point below adequate.

Table 4: Year 1 *READ 180* Fidelity of Implementation Teacher/Classroom Ratings

Note: ID numbers were randomly generated and hold no meaning

Teacher/ Classroom ID	Prof. Dev. Rating	Observation(s) Ratings			Teacher Survey Ratings	SAM Ratings	Overall Rating
		Feb.	May	Avg.			
910	3	3	3	3.0	3.8	4	3.45
707	4	3	4	3.5	3.0	3	3.38
899	4	2	4	3.0	3.2	3	3.30
821	2	4	4	4.0	3.0	4	3.25
322	3	4	3	3.5	3.8	2	3.08
397	3	3	4	3.5	3.8	2	3.08
604	4	3	*	3.0	3.2	2	3.05
848	3	4	2	3.0	3.0	3	3.00
628	3	2	*	2.0	3.8	3	2.95
122	3	3	*	3.0	3.0	2	2.75
221	3	3	2	2.5	2.8	2	2.58
694	3	2	*	2.0	3.2	2	2.55
513	3	2	3	2.5	*	2	2.50
727	1	*	3	3.0	*	3	2.33
242	1	2	*	2.0	2.0	3	2.00
298	2	3	3	3.0	*	1	2.00
516	3	2	*	2.0	*	1	2.00
380	2	3	1	2.0	2.6	1	1.90
895	0	3	*	3.0	*	1	1.67
Number “Adequate”	13	12	9	12	11	8	8
Percent “Adequate”	68.4	66.7	75.0	66.7	78.6	42.1	42.1

* Data were not gathered or were not available.

Year 2 Implementation Study

Professional Development Levels

In Year 2, the professional development participation score was again developed by combining the number of points assigned to different types of professional development. However, the online course was included for the Year 2 ratings, and the scoring was adjusted. There were four types of professional development included in this calculation:

- 1.attendance at each (of two) all-day session earned a 2,
- 2.attendance at each (of four) networking meeting earned a 1
- 3.the first year of experience teaching *READ 180* earned a 2, and any number of years of experience beyond that earned a 3
- 4.completion of the Scholastic’s online course on or before October 31, 2007, earned a 3, and completion by January 31, 2008, earned a 1.5.

This resulted in a possible total of 12 points. Evaluators determined that professional development scale scores greater than or equal to ten would be considered “high,” those seven or higher would be considered “moderate,” those four or higher would be “low,” and scores three and below would be “minimal.” These ratings are included in the second column of Table 5. In order to receive a “high” rating, a teacher would have to participate in at least three types of professional development, including completion of at least half of the online course.

Levels and Variability of Implementation at the Classroom Level

Implementation fidelity and variability were monitored through classroom observations, student surveys, and data generated by the Scholastic Achievement Manager. In Year 2, as in Year 1, the variables used focused on what was made available to or completed by students, not how well students performed on tasks or assessments. For example, the student survey administered in the spring of 2008 asked students how many books they read in *READ 180* class during the past year, but not how well they did on the quizzes they took.

Evaluators first examined all included variables and created an equation for translating each survey, SAM, or observational variable to the four-point scale. Second, an equation was created that encompassed the data from within each source of data (surveys, SAM, observations, and professional development). Finally, those scores were averaged to create the ratings that appear in Table 5. All of the 19 classrooms had SAM and survey data; each classroom was observed at least four times.

Whole-group instruction and independent reading are activities that, according to interviews with school staff members, occur in many regular language arts classes. Therefore, data related to the small-group instruction and computer rotations were weighed more heavily because they are components of the *READ 180* program that distinguish the program from the regular language arts classes offered. Also, in the calculations of overall classroom implementation level, observations were weighed the most heavily because they addressed all components of *READ 180*. Student surveys were weighted the second most heavily because they addressed two components: small-group instruction and independent reading. These were followed by SAM

data and the professional development scale. (The equations used to development the ratings are included in Appendix M.)

During the 2007–2008 school year, members of the evaluation team completed classroom observations (N =44) during September (n=17), February (n=17), and May (n=10). MCS staff members completed observations (N =37) during October (n=19) and April (n=18), and Scholastic representatives completed observations (N =18) during December. The evaluation team used a different observation protocol than the MCS and Scholastic observers. Therefore, evaluators identified 15 items that were similar between the two protocols and used those to calculate the observation ratings. The 15 items used related to the following:

- the timing of the class (that is, whether the class had four 20-minute and one 10-minute segments)
- the number of students (seven or fewer) in each small group
- the layout of the room
- content and student engagement in whole-group instruction and the small-group instruction, computer, and independent reading rotations
- use of the Scholastic *flexBook* and/or other *READ 180* materials for instruction

Student surveys, which were based on a *READ 180* survey developed by Scholastic, were administered by MCS during the spring of 2008. Survey items used for calculating the classroom implementation rating asked about the number of books students read during the independent reading rotation and the “workshops” (or chapters) that the students read as part of their class. Workshops are generally related to small-group instruction, and this item, when averaged across all of a classroom’s students, indicated how much of the textbook was covered.

SAM data are automatically generated for all students in *READ 180*, so evaluators linked all students with their teachers and averaged the data from all students taught by each teacher. The SAM variables used for this report included the following:

- average number of *READ 180* computer software sessions per week
- average daily number of minutes spent in these sessions

Table 5: Year 2 *READ 180* Fidelity of Implementation Teacher/Classroom Ratings*Note: ID numbers were randomly generated and hold no meaning*

Classroom ID	Professional Development	Observation Average	Student Surveys	SAM	Overall Rating
3566	3.0	3.70	3.5	3.5	3.63
8348	3.0	3.71	3.5	3.0	3.45
5224	2.0	3.52	3.5	2.5	3.43
3328	2.0	3.41	3.5	3.5	3.39
6132	4.0	3.59	3.5	2.5	3.36
5541	3.0	3.49	3.5	3.5	3.31
3801	3.0	3.14	3.5	3.0	3.30
5546	3.0	3.58	3.5	2.5	3.28
2026	4.0	3.53	3.0	2.5	3.21
6033	4.0	3.38	3.5	3.0	3.11
4781	2.0	3.25	3.0	2.5	3.10
4420	2.0	3.36	2.0	3.5	3.09
3973	3.0	2.75	3.0	3.0	2.95
6410	3.0	3.23	3.0	3.0	2.94
2918	3.0	2.94	3.0	2.5	2.90
2109	4.0	3.13	3.0	2.5	2.78
5499	3.0	3.47	1.5	2.5	2.54
5515	3.0	2.42	3.0	2.0	2.47
8877	2.0	2.38	2.0	3.0	2.35
Number "Adequate"	14	15	16	10	12
Percent "Adequate"	74%	78%	84%	53%	63%

Conclusions Regarding Implementation of the Targeted Intervention

An analysis of the *READ 180* program data gathered and obtained found wide variation in implementation across classrooms in the eight schools. The data show that implementation improved from Year 1 to Year 2. Table 6 shows the numbers of classrooms, out of 19 in both years, that were rated as adequate in Years 1 and 2. Several factors caution readers against drawing inferences related to the comparison between the ratings from Years 1 and 2. First, the amount of data gathered and used for the Year 2 classroom ratings was larger: the number of observations tripled and student survey was made available. Second, there were changes in the way that professional development was offered. Third, many classrooms were not equipped with working *READ 180* software and computers until the late fall of Year 1 (which would decrease ratings based on SAM data). Also, seven teachers left after the end of Year 1 and were replaced in Year 2. These changes prevent within-classroom comparisons between Year 1 and Year 2.

Despite these cautions, it does appear that more classrooms had higher levels of implementation in Year 2. The thoroughness of the implementation data that are gathered and subsequent ratings and analyses will continue to improve through Years 3 and 4. As noted on page 50, further analyses will be conducted to explore the correlations between these ratings and impact

findings at the classroom level and between the district’s student attendance data and impact findings.

Table 6: Numbers of teachers Rated Adequate in Years 1 and 2

Program Component	Year 1	Year 2	Change
Professional development	13	14	+1
Observations	12	15	+3
Teacher surveys	11*		n/a
Student surveys		16	n/a
SAM	8	10	+2
Overall	8	12	+4

* This is 11 out of the 14 surveys that were returned, or approximately 78 percent.

Description of the Counterfactual for the Targeted Intervention

The targeted intervention is supplemental, so all students identified as the target population should also have been enrolled in a “regular” language arts class, whether or not they were selected to be enrolled in *READ 180*. Treatment students in three schools receive a truncated period for language arts: they are enrolled in a two-hour class that combines 90 minutes of *READ 180* with 30 minutes of language arts instruction, while control students in these schools are enrolled in “regular” language arts classes that last 45 to 55 minutes.

Students in both the treatment and control groups might also participate in additional classes related to language arts. Some of these are reading classes, and this is made clear in the class name. Table 7 provides a list of the classes offered by each of the MSRP schools, sorted by grade. Class registration data were collected to detail how many treatment and control students were assigned to these classes. However, the class registration data (which are different from the school enrollment data used for the impact analyses) contain discrepancies. The evaluation team is currently working with MCS staff to determine whether these data can be clarified or corrected. If this is possible, analyses of these better data will be included in the exploratory analyses and in subsequent implementation reports. (The team is also exploring whether there is an actual discrepancy between the name of some of the classes and the number of days those classes are said to last, e.g., “Reading 7-180,” which is identified as lasting 90 days.)

Table 7: ELA-Related Courses Offered at MSRP schools

Course Name	Grade	# Days	A. Maceo Walker	American Way	Corry	Hamilton	Hickory	Lanier	Riverview	Sherwood
Comm Skills 4-6	6	180		X						
Creative Writing 6-180	6	90							X	
Creative Writing 6-90	6	90					X			
English Skills 4-6	6	180	X			X		X		X
Language Arts 06	6	180	X	X	X	X	X	X	X	X
READ 180	6	180	X	X	X	X	X	X	X	X
Reading 6-180	6	180	X	X	X	X	X			X
Reading Skills 4-6	6	180		X		X				X
Writing 6-90	6	90								X
Content Area Reading 6-8	6, 7, 8	90					X			
Word Bldg Exp 6-8/180	6, 7, 8	180					X			
Creative Writing 7-90	7	90		X			X		X	X
Language Arts 07	7	180	X	X	X	X	X	X	X	X
Language Arts 07 Honors	7	180	X	X	X	X	X			
READ 180	7	180	X	X	X	X	X	X	X	X
Reading 7-180	7	180	X		X		X		X	X
Reading 7-90	7	90		X						
Reading 7-90	7	180						X		
Fail Free Read Lab 7-8/180	7, 8	180			X					
Comm Skills 09	8	180				X	X			
Creative Writing 8-180	8	90							X	
Creative Writing 8-180	8	180	X						X	
Creative Writing 8-90	8	90		X			X			X
Language Arts 08	8	180	X	X	X	X	X	X	X	X
Language Arts 08 Honors	8	180	X	X	X	X	X			
READ 180	8	180	X	X	X	X	X	X	X	X
Reading 8-180	8	180			X		X			X
Reading 8-90	8	180						X		
Word Bldg Expl 8-90	8	90		X			X			

IV. Evaluation of the Impacts of the Targeted Intervention: Years 1 and 2

Study Design

This section describes the design of the evaluation of the impacts in Years 1 and 2 of *READ 180*. It includes explanations of how the student samples were constructed and when and how data were collected and analyzed.

Sample Selection

At the beginning of Year 1, students who were struggling readers were identified in the sixth to eighth grades in each participating school. The intent was to randomly assign approximately 40 struggling readers in each grade in each school to receive *READ 180* services. Prior to random assignment, the participating schools were given the opportunity to opt out students from being eligible (e.g., due to parent objections, assignment to self-contained classrooms for special education services, or teacher judgments that TCAP scores were not representative of students' higher achievement levels).⁶

Struggling readers who were not opted out and who had not received *READ 180* services in the previous two school years were deemed eligible for random assignment to *READ 180*. Within each grade at each school, 40 students were randomly selected from this eligible pool to receive *READ 180* services. These 40 students were assigned by the school to one of two or three *READ 180* classes in each grade. The students who were not randomly selected to receive *READ 180* services were assigned to the control group.⁷ Also, due to significant variations in the way each school would or would not permit students receiving special education services to be assigned to *READ 180*, it was decided to base the *READ 180* impact analyses on treatment and control students not receiving special education services. The remaining students constitute the intent-to-treat (ITT) group for the analyses of *READ 180* impact on student outcomes in Year 1. Table 8 describes the number of treatment, control, and non-eligible students enrolled in the eight MSRP schools at the time of random assignment (September 18, 2006).

At the beginning of the 2007–2008 school year (Year 2), struggling readers were identified in sixth grade in each participating school. As in Year 1, the intent was to randomly assign approximately 40 struggling readers in sixth grade in each school to receive *READ 180* services. Prior to random assignment, the participating schools were given the opportunity to opt out students from being eligible (e.g., due to parent objections, assignment to self-contained classrooms for special education services, or teacher judgments that TCAP scores were not representative of student achievement levels). Students who had received *READ 180* services during the previous two years also were excluded from the eligible pool.

⁶ There were a few treatment group students that were opted out after random assignment. Since the control group counterparts for those students could not be identified, the treatment group students that were opted out after random assignment were retained in the treatment group for the intent to treat (ITT) analyses.

⁷ Some *READ 180* teachers also taught regular English/Language Arts classes; therefore, the evaluators examined the possibility of cross-contamination across control and treatment groups. Five *READ 180* teachers were identified who taught a total of 35 control students in regular ELA classes. Of these 35 students, 31 were enrolled in a single school. Evaluators will continue to monitor this possible path of “cross-contamination” for the remainder of the project.

Table 8: Number of Students Enrolled in Striving Reader Study Schools in Year 1 By *READ 180* Design Group

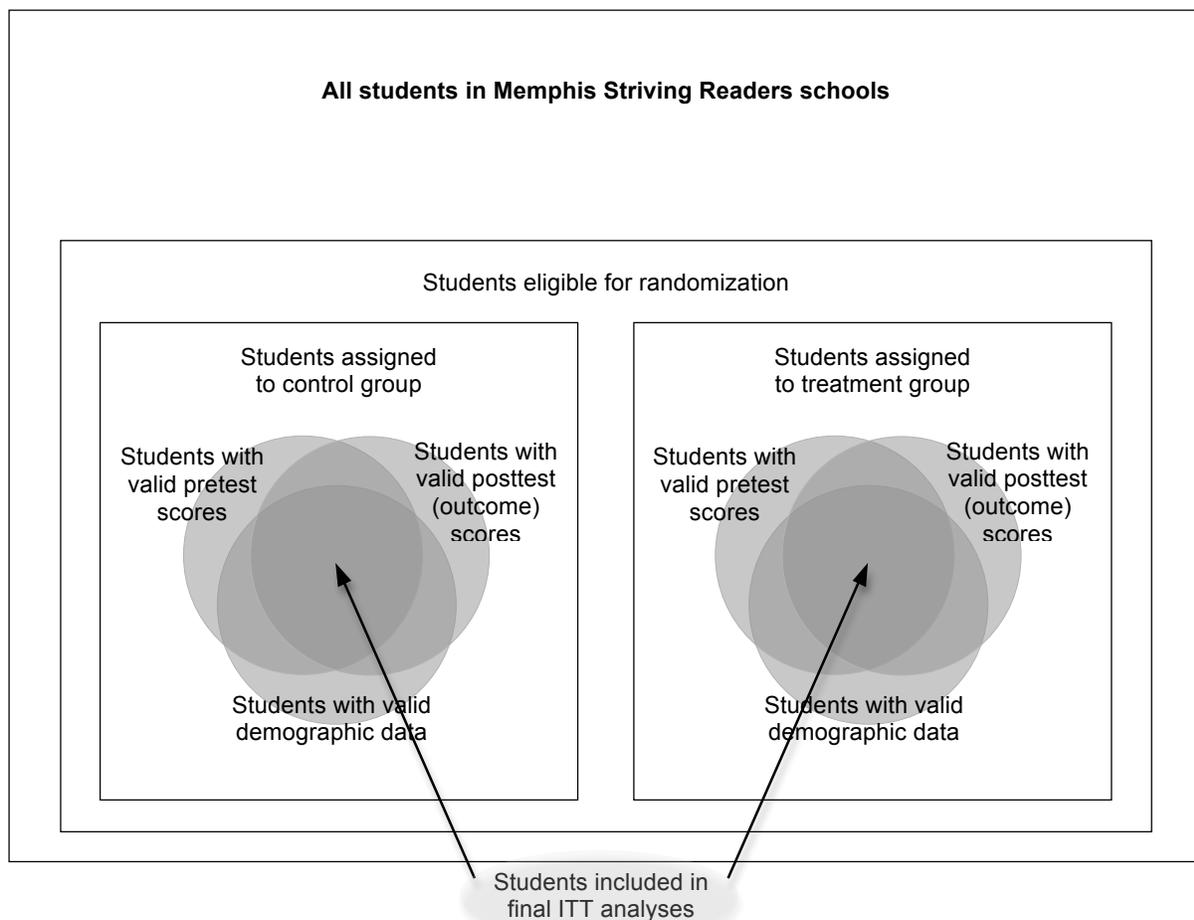
Grade	<i>READ 180</i> Treatment Group ^b	<i>READ 180</i> Control Group ^b	Non-Eligible Students	Total MSRP School Enrollment ^a
6	239	392	903	1540
7	233	370	1270	1880
8	226	280	1253	1767
All	698	1042	3426	5187

^a Enrollment as of 9/18/06

^b These two groups constitute the ITT group (N=1740) for *READ 180* immediate impact analyses in Year 1.

Tables eight through ten and tables 11 through 21 all provide information about students in the Striving Readers schools. The largest numbers, such as those included in tables eight and nine, include all students enrolled in Striving Readers schools on the date provided. The numbers of students included in the following tables are increasingly smaller and sometimes vary by outcome measure. This is because of the steps taken to determine whether students can be included in the final ITT analyses. Figure 5 illustrates the decreasing numbers of students and the path taken to identify students included in the final analyses. Details about the reasons for attrition of students in the ITT analyses are provided in the tables included as Appendix M-6.

Figure 5: Illustration of student samples used for ITT analyses



Within sixth grade at each school, between 30 and 40 students were randomly selected from this eligible pool to receive *READ 180* services.⁸ These students were assigned by the school to one of two or three *READ 180* classes in each grade. The students who were not randomly selected to receive *READ 180* services were assigned to the control group. These sixth-grade students constitute the intent to treat (ITT) group for the analyses of the immediate impact of *READ 180* on student outcomes in Year 2. (As in Year 1, students receiving special education services were excluded from the impact analyses.) Table 9 shows the number of treatment, control, and non-eligible sixth-grade students enrolled in the eight MSRP schools at the time of random assignment in Year 2 (September 7, 2007).

Table 9: Number of Grade 6 Students Enrolled in Striving Reader Study Schools in Year 2 by *READ 180* Design Group

Grade	<i>READ 180</i> Treatment Group ^b	<i>READ 180</i> Control Group ^b	Non-Eligible Students	Total Striving Reader School Enrollment ^a
6	289	404	734	1427

^a Enrollment as of 9/7/07

^b These two groups constitute the ITT group (N=693) for *READ 180* immediate impact analyses in Year 2.

In order to evaluate the long-term (two-year) impact of *READ 180* services, students from the Year 1 ITT analysis group who were enrolled in the seventh and eighth grades in a participating school when the ITBS was administered in the spring of Year 2 were identified.⁹ Table 10 describes the number of these students compared to the original number assigned to *READ 180* treatment and control groups in Year 1. The amount of attrition is relatively constant across experimental groups and grades.

Table 10: Number of Students Remaining in *READ 180* ITT Group in Spring of Year 2 Compared to Number of Students in Year 1 by Grade and Design Group

Grade		Design Group	Year 1	Year 2	Attrition Rate
During Year 1	During Year 2				
6	7	<i>READ 180</i>	239	160	33%
		Control	392	260	34%
7	8	<i>READ 180</i>	233	159	32%
		Control	370	241	35%
Both Grades		<i>READ 180</i>	472	319	32%
		Control	762	501	34%
		Both Groups	1,234	820	34%

⁸ Due to declining enrollments in several of the schools, it was not possible to assign 40 students to *READ 180* and still have a reasonable number of control students (at least half the number of *READ 180* students).

⁹ This identification process was consistent with an earlier procedural decision to include in the ITT analyses only those students who were enrolled in participating schools at the time of the spring administration of the ITBS.

Data Collection

As noted on page 15, the measures of student outcomes for determining the impact of *READ 180* on struggling readers are the ITBS and the TCAP. Standard scores were used to measure reading comprehension, vocabulary, and total reading on the ITBS. These scores are vertically equated across grade levels so that students in higher grades achieve higher scores on average.¹⁰ The ITBS was administered twice in Year 1—during the week beginning September 18, 2006, and during the week beginning April 30, 2007—by classroom teachers to all students in the MSRP schools, except those in self-contained special education classrooms and a very small number whose parents did not consent to the student’s participation in the testing. The spring 2007 test scores measured treatment and control student reading achievement levels at the end of Year 1. The fall 2006 test scores were used to control for random differences in reading achievement levels between treatment and control students at the beginning of the year, as well as reduce the within-school error variance in the spring 2007 test scores.

The ITBS was also administered twice in Year 2—during the weeks beginning September 17, 2007, and May 12, 2008—by classroom teachers in MSRP schools. The fall administration was only for students in sixth grade; all students in MSRP schools took the spring administration. The spring 2008 test scores measured treatment and control student reading achievement levels at the end of Year 2. The fall 2007 test scores were used as control variables for cross-sectional, immediate impact analyses of Year 2 achievement for sixth-grade students. The fall 2006 test scores were used as control variables for cross-sectional, long-term impact analyses of Year 2 achievement and as baseline test scores for longitudinal analyses of growth in achievement over the two years for students in the seventh and eighth grades in Year 2.

The TCAP is administered by MCS for the state on or about the first week in April each year. The spring 2007 test scores measured treatment and control student achievement levels in the four core content areas at the end of Year 1. The spring 2006 scores in the same content area were used to control for random treatment-control differences and reduce within-school error variance in spring 2007 scores. The spring 2007 test scores were used as control variables for cross-sectional, immediate impact analyses of Year 2 achievement for sixth-grade students. The spring 2006 test scores were used as control variables for cross-sectional, long-term impact analyses of Year 2 achievement and as baseline test scores for longitudinal analyses of growth in achievement over the two years for students in the seventh and eighth grades in Year 2.

Data Analysis

Cross-sectional ITT impact analyses of student achievement in reading and the four core content areas were conducted to assess the immediate effects of first-year participation in *READ 180* on student outcomes for the 1,740 eligible struggling readers in the sixth, seventh, and eighth grades in Year 1. Also, to investigate any suggestions of interactions of *READ 180* impact and grade level, separate analyses were conducted for students in the sixth, seventh, and eighth grades. A separate cross-sectional analysis of the 693 eligible students in sixth grade in Year 2

¹⁰ Researchers tried to obtain the psychometric properties for the TCAP, but these are not available. The ITBS is a nationally recognized measure with adequate psychometric credentials.

was conducted to determine whether the immediate impact of *READ 180* varied between Years 1 and 2.

Similar cross-sectional analyses were conducted to assess the long-term effects of participating in *READ 180* for two years on student achievement at the end of Year 2 for the remaining 820 ITT students in the seventh and eighth grades. In addition, the long-term effects of *READ 180* on growth in student achievement during Years 1 and 2 were examined using longitudinal analyses of the achievement of these 820 students at the beginning and end of Year 1 and the end of Year 2.

Multi-level regression analysis models were used to estimate and test the statistical significance of the difference between the achievement of students receiving *READ 180* and the control group. Two-level models were employed for the Year 1 cross-sectional analyses that express the spring ITBS and TCAP scores as a function of student and school variables.¹¹ The spring 2007 ITBS and TCAP scores were the dependent variables. The 2006 ITBS and TCAP scores—representing the same test or subject as the dependent variable—were included as the principal student-level covariate. Other control variables at the student and school level were tested for inclusion as covariates in these analyses. The *READ 180* treatment variable was included at the student level of these models.

Similar two-level models were employed for the cross-sectional analyses of student achievement at the end of Year 2. The only differences were the use of spring 2008 ITBS and TCAP test scores as the dependent variables and, for the sixth-grade students, the use of 2007 ITBS and TCAP test scores as one of the student covariates.

Three-level, longitudinal models were used to estimate and test the statistical significance of the difference between growth in achievement over the two years for students receiving *READ 180* and the control group. The first level expresses each student's test scores at three points in time as the sum of the baseline score plus the growth at the end of Year 1 and the additional growth at the end of Year 2. The second level expresses the student's baseline test score as a function of receiving *READ 180* and other student characteristics, and expresses the Year 1 and Year 2 growth in test scores as a function of receiving *READ 180*. The third level expresses the average baseline score as a function of school covariates. The complete specification of the multi-level regression models employed to determine the immediate and long-term impacts of the *READ 180* intervention is provided in Appendix A.

Table 11 summarizes the dependent and independent variables and the covariates included in these analyses.

¹¹ Three-level models, employing school, teacher, and student variables were explored. These analyses proved to be relatively complex and equivocal due to each student's having different teachers for the core content areas and significant amounts of missing teacher data. Also, the results did not vary noticeably from the results of the two-level models. The evaluation team decided to omit these models from the impact analyses.

Table 11: All Variables Included in READ 180 Impact Analytical Models for Years 1 and 2

Variable	Level	Coding / Range
Dependent		
Spring 2007/2008 ITBS Total Reading*	Student	Standard Score 100-350
Spring 2007/2008 ITBS Comprehension*	Student	Standard Score 100-350
Spring 2007/2008 ITBS Vocabulary*	Student	Standard Score 100-350
Spring 2007/2008 TCAP Reading/LA*	Student	Scale Score 300-750
Spring 2007/2008 TCAP Mathematics*	Student	Scale Score 300-750
Spring 2007/2008 TCAP Science*	Student	Scale Score 100-300
Spring 2007/2008 TCAP Social Studies*	Student	Scale Score 100-300
Independent		
READ 180 Participation	Student	Yes = 1; No = 0
Covariates		
Test Score at End of Year 1	Time	Yes = 1; No = 0
Test Score at End of Year 2	Time	Yes = 1; No = 0
Fall 2006/2007 ITBS Total Reading** ***	Student	Standard Score 100-350
Fall 2006/2007 ITBS Comprehension** ***	Student	Standard Score 100-350
Fall 2006/2007 ITBS Vocabulary** ***	Student	Standard Score 100-350
Spring 2006/2007 TCAP Reading/LA** ***	Student	Scale Score 300-750
Spring 2006/2007 TCAP Mathematics** ***	Student	Scale Score 300-750
Spring 2006/2007 TCAP Science** ***	Student	Scale Score 100-300
Spring 2006/2007 TCAP Social Studies** ***	Student	Scale Score 100-300
Gender	Student	Female = 1; Male = 0
African-American	Student	Yes = 1; No = 0
Hispanic	Student	Yes = 1; No = 0
Free/Reduced Lunch (Fall 2006/2007)***	Student	Yes = 1; No = 0
English Language Learner (Fall 2006/2007)***	Student	Yes = 1; No = 0
Enrolled in Grade 7 in Year 1	Student	Yes = 1; No = 0
Enrolled in Grade 8 in Year 1/Year 2*	Student	Yes = 1; No = 0
Percentage Female (Fall 2006)	School	0-100
Percentage African-American (Fall 2006)	School	0-100
Percentage Special Ed (Fall 2006)	School	0-100
Percentage FRL (Fall 2006)	School	0-100
Percentage ELL [†] (Fall 2006)	School	0-100
School Enrollment (Fall 2006)	School	400-1200

* Second date applies for analyses in Year 2

**Only used as a covariate in cross-sectional, not in longitudinal, analyses

*** Second date applies for Grade 6 analyses in Year 2

[†] English language learners

Selection of Covariates

There are different approaches to including and/or excluding covariates in multi-level regression, as there are in single-level regression analyses. The approach that was used in these analy-

ses was to (1) include all student- and school-level covariates in the model,¹² (2) run the model, (3) eliminate the school covariate with the lowest significance level (highest p-value) not less than 0.2, (4) repeat steps 2 and 3 until the remaining covariates had p-values less than 0.2, and (5) repeat steps 2 to 4 for the student covariates.

Treatment of Missing Data

The only variables in these analyses that contained missing data were the ITBS and TCAP standard and scale scores. The number of students for whom test scores were missing in any analysis was small relative to the total number of students such that the power of these analyses would not be significantly compromised by simply omitting these students. Thus, procedures for imputing missing values were not employed. The amount of attrition due to missing test scores and possible differential attrition between treatment and control groups were studied. The results of this study are presented in the discussion of the results of the *READ 180* impact analyses below. (See Appendix M for a summary table that describes missing data in greater detail).

Description of the Year 1 and Year 2 Samples of Students for *READ 180* Impact Analyses

Equivalence on Student Demographic Characteristics

As described above, 1,740 eligible struggling readers were randomly assigned to the *READ 180* treatment or control groups in Year 1 of the MSRP study. The grade level and other demographic characteristics of these students are presented in Table 12.

The number of students decreases as the enrollment grade increases, and this difference is reflected more strongly in the control group since approximately equal numbers were randomly assigned to the treatment group in each grade. These differences in grade enrollment between treatment and control groups emphasize the importance of treating the student's enrollment grade as a covariate in the analyses of *READ 180* impact. Also, all but two students were either African-American or Hispanic, which supported the creation of two dichotomous covariates to represent membership in these two race/ethnicity groups. Finally, the differences in demographic composition of the treatment and control groups were relatively minor, although some were statistically significant given the large number of students overall. Including these characteristics as student-level covariates in the analytical models helps to control for these small differences, as well as reduce the within-school error variance in the dependent variables.

¹² In the longitudinal models, the two variables representing the time at which the test score was obtained are always included in the first level of the model.

Table 12: Demographic Characteristics of the Year 1 *READ 180* ITT Sample

Student Characteristic	Control ^a	<i>READ 180</i> ^a	Total ^a
Enrolled in Grade 6	392 (37.6%)	239 (34.2%)	631 (36.3%)
Enrolled in Grade 7	370 (35.5%)	233 (33.4%)	603 (34.7%)
Enrolled in Grade 8	280 (26.9%)	226 (32.4%)	506 (29.1%)
Female	465 (44.6%)	286 (41.0%)	751 (43.2%)
Male	577 (55.4%)	412 (59.0%)	989 (56.8%)
African-American	955 (91.6%)	657 (94.1%)	1612 (92.6%)
Hispanic	86 (8.2%)	40 (5.7%)	126 (7.2%)
Free or Reduced Lunch	931 (89.3%)	619 (88.7%)	1550 (89.1%)
English Language Learner	83 (8.0%)	34 (4.9%)	117 (6.7%)
Total	1042	698	1740

^a Percentages are based on the total number of students in the control, treatment, or total group.

In Year 2, 693 eligible struggling readers in sixth grade were randomly assigned to the *READ 180* treatment or control groups. The demographic characteristics of these students are presented in Table 13. Other than gender, where the control group has significantly more males, the treatment and control groups are quite similar.

Table 13: Demographic Characteristics of the Year 2 Grade 6 *READ 180* ITT Sample

Student Characteristic	Control ^a	<i>READ 180</i> ^a	Total ^a
Female	169 (41.8%)	143 (49.5%)	312 (45.0%)
Male	235 (58.2%)	146 (50.5%)	381 (55.0%)
African-American	382 (94.6%)	274 (94.8%)	656 (94.7%)
Hispanic	21 (5.2%)	15 (5.2%)	36 (5.2%)
Free or Reduced Lunch	382 (94.6%)	276 (95.5%)	658 (94.9%)
English Language Learner	17 (4.2%)	14 (4.8%)	31 (4.5%)
Total	404	289	693

^a Percentages are based on the total number of students in the control, treatment, or total group.

Also in Year 2, 820 students in the seventh and eighth grades remained enrolled in a participating school out of the 1,234 eligible struggling readers in the sixth and seventh grades in the Year 1 *READ 180* ITT group. The demographic characteristics of these “stayers,” compared with the 414 “leavers” who were not in a participating school on May 13, 2008, are presented in Table 14.

As noted in Table 14, differences between treatment and control groups for the 820 seventh- and eighth-grade students in Year 2 who remained in the *READ 180* ITT sample are relatively small. There are higher percentages of males, African-Americans, and students receiving free or reduced-price lunch in the treatment group and higher percentages of Hispanic and English language learner (ELL) students in the control group. Also, the differences between treatment and control groups for the “stayers” appear similar to the differences between treatment and control groups for the 414 “leavers,” with two exceptions. The difference between treatment and control percentages of male students was larger for the 414 “leavers,” and the percentage of treatment

students receiving free or reduced-price lunch was a few points higher than it was for the controls in the “stayers,” while it was a few points lower for treatment students in the “leavers.” Overall, however, differences in treatment and control groups are very similar for those who remained and those who did not.

Table 14: Demographic Characteristics of the Year 2 “Stayers” and “Leavers” from the Year 1 *READ 180* ITT Sample

	“Stayers”		“Leavers”	
	Control ^a	<i>READ 180</i> ^a	Control ^a	<i>READ 180</i> ^a
Enrolled in Grade 6 → 7	260 (51.9%)	160 (50.2%)	132 (50.6%)	79 (51.6%)
Enrolled in Grade 7 → 8	241 (48.1%)	159 (49.8%)	129 (49.4%)	74 (48.4%)
Female	236 (47.1%)	139 (43.6%)	118 (45.2%)	61 (39.9%)
Male	265 (52.9%)	180 (56.4%)	143 (54.8%)	92 (60.1%)
African-American	457 (91.2%)	297 (93.1%)	238 (91.2%)	144 (94.1%)
Hispanic	43 (8.6%)	21 (6.6%)	22 (8.4%)	9 (5.9%)
Free or Reduced Lunch	440 (87.8%)	287 (90.0%)	240 (92.0%)	137 (89.5%)
English Language Learner	46 (9.2%)	18 (5.6%)	19 (7.3%)	7 (4.6%)
Total	501	319	261	153

^a Percentages are based on the total for the control and treatment groups for each type of student.

Equivalence on Baseline Achievement

Comparisons between treatment and control groups on the baseline 2006 ITBS and TCAP test scores were carried out for the 1,740 students in the Year 1 ITT sample and the 820 Year 2 “stayers.” Treatment and control comparisons on the baseline 2007 ITBS and TCAP test scores were carried out for the Year 2 sixth-grade ITT sample.

Table 15 describes the differences between Year 1 *READ 180* treatment and control groups on baseline 2006 test scores for the three ITBS standard scores and the four TCAP content area assessments. With random assignment, the treatment and control groups should be very similar on all seven test scores.

The treatment group performed higher on all seven test scores. However, the significance level for each estimated difference is greater than 0.05, allowing one to conclude that the treatment and control groups are statistically equivalent in terms of their baseline 2006 ITBS and TCAP test scores.

Table 15: Comparison of Year 1 *READ 180* Treatment and Control Groups on Baseline 2006 Scores on Each Achievement Test

Test Score	Means		Signif. Level
	Control	READ 180	
ITBS Total Reading Standard Score	190.2 (940) ^a	191.5 (656)	0.097
ITBS Comprehension Standard Score	186.2 (944)	188.0 (660)	0.059
ITBS Vocabulary Standard Score	194.1 (950)	195.1 (658)	0.354
TCAP Reading/LA Scale Score	474.8 (1042)	476.5 (698)	0.188
TCAP Mathematics Scale Score	484.4 (1040)	487.4 (697)	0.062
TCAP Science Scale Score	179.1 (1006)	180.5 (686)	0.116
TCAP Social Studies Scale Score	184.0 (1007)	184.5 (685)	0.593

^a Numbers in parentheses are the numbers of students in each group having a valid test score.

Table 16 describes the differences between Year 2 sixth-grade *READ 180* treatment and control groups on baseline 2007 test scores. Again, with random assignment, the 693 students in the treatment and control groups should be very similar on all seven test scores.

Table 16: Comparison of Year 2 Grade 6 *READ 180* Treatment and Control Groups on Baseline 2007 Scores on Each Achievement Test

Test Score	Means		Signif. Level
	Control	READ 180	
ITBS Total Reading Standard Score	184.9 (364) ^a	182.0 (247)	0.007
ITBS Comprehension Standard Score	182.8 (365)	179.7 (247)	0.014
ITBS Vocabulary Standard Score	186.9 (365)	184.7 (252)	0.104
TCAP Reading/LA Scale Score	487.8 (403)	482.4 (289)	0.004
TCAP Mathematics Scale Score	489.8 (403)	487.3 (289)	0.208
TCAP Science Scale Score	184.4 (400)	182.8 (283)	0.204
TCAP Social Studies Scale Score	189.5 (399)	188.9 (283)	0.551

^a Numbers in parentheses are the numbers of students in each group having a valid test score.

The control group performed higher on all seven test scores. As shown in Table 16, this advantage was statistically significant for the ITBS Comprehension and Total Reading standard

scores and for the TCAP Reading/LA scale score. Thus, even though students were assigned to treatment and control groups randomly, the control group scored significantly higher on most of the baseline reading measures. Treating the 2007 test scores as covariates in the analyses of the impact of *READ 180* on 2008 test scores allows some adjustment to be made for these differences.

Table 17 describes the 2006 baseline test score differences between the 820 treatment and control students from the Year 1 *READ 180* ITT sample who “stayed” in a participating school in Year 2 in seventh and eighth grade.

Table 17: Comparison of Year 2 “Stayers” from *READ 180* Treatment and Control Groups on Baseline 2006 Scores on Each Achievement Test

Test Score	Means		Signif. Level
	Control	READ 180	
ITBS Total Reading Standard Score	186.7 (472) ^a	188.1 (305)	0.217
ITBS Comprehension Standard Score	183.0 (474)	185.1 (305)	0.096
ITBS Vocabulary Standard Score	190.4 (476)	190.0 (306)	0.674
TCAP Reading/LA Scale Score	472.7 (501)	471.6 (319)	0.558
TCAP Mathematics Scale Score	484.2 (501)	487.0 (318)	0.183
TCAP Science Scale Score	180.9 (482)	182.1 (315)	0.290
TCAP Social Studies Scale Score	185.3 (482)	186.1 (316)	0.542

^a Numbers in parentheses are the numbers of students in each group having a valid test score.

These treatment-control differences are all small and not significant (as shown in Table 17). They also appear very similar to, if slightly lower than, the treatment-control differences for the entire Year 1 ITT sample. Two-way ANOVAs were used to test for an interactive effect of “staying” and treatment/control group membership on 2006 test scores. No significant interactions were found.

Impact of *READ 180* Participation on Student Achievement in Years 1 and 2

Immediate Impact of *READ 180* in Year 1

The multi-level regression model described in Appendix A was used to estimate the immediate impact of *READ 180* on the reading and content area achievement of the 1,740 eligible struggling readers in the Year 1 *READ 180* ITT sample. Table 18 summarizes the results of these analyses for student reading achievement measured by the ITBS and student achievement in the four core content areas measured by the TCAP. (The complete results of the multi-level analyses of the *READ 180* impact on these seven test scores can be found in Appendix B in Tables B1–

B7.) The table displays several statistical parameters. The **unadjusted means** show the actual mean 2007 test scores for the treatment and control groups. The numbers in parentheses at the bottom of these cells is the number of students in the respective group with a valid test score. The **adjusted means** are the average scores controlling for all covariates retained in the analytical model—the variable indicating treatment/control group membership and all “significant” covariates ($p < 0.2$). The **estimated impact** is the difference between the treatment and control group adjusted means (treatment minus control). A positive impact means the *READ 180* treatment group averaged higher achievement on the particular test than the control group, controlling for covariates included in the final analytical model (see Appendix B, Tables B1–B7). A negative impact means the control group averaged higher than the treatment group. The **significance level** and **effect size** are two indicators of the importance of the estimated difference. Conventionally, a significance level less than 0.05 is an acceptable indication that the estimated difference is not due to chance, i.e., that it is “statistically significant.” Conventionally, an effect size between 0.2 and 0.5 is considered small, between 0.5 and 0.8 is medium, and greater than 0.8 is large.

Table 18: Immediate Year 1 Impact of *READ 180* on Spring 2007 Scores on Each Achievement Test

Test Score	Unadjusted Means		Adjusted Means		Est. Impact	Effect Size ^a	Signif. Level
	Control	READ 180	Control	READ 180			
ITBS Total Reading Standard Score	191.8 (712) ^b	192.9 (511)	192.6	192.1	-0.5	0.03	0.532
ITBS Comprehension Standard Score	186.7 (718)	187.6 (519)	187.0	187.0	0.0	0.00	0.976
ITBS Vocabulary Standard Score	197.0 (726)	198.3 (519)	197.5	197.6	0.1	0.01	0.937
TCAP Reading/LA Scale Score	495.8 (972)	498.0 (664)	496.9	497.1	0.2	0.01	0.882
TCAP Mathematics Scale Score	500.0 (971)	501.8 (661)	500.0	500.2	0.2	0.01	0.904
TCAP Science Scale Score	185.1 (915)	185.6 (643)	185.6	185.1	-0.5	0.03	0.573
TCAP Social Studies Scale Score	185.1 (906)	186.1 (644)	185.0	185.8	0.8	0.05	0.323

^a The method used to calculate effect size was Glass's Δ , the difference between treatment and control groups' adjusted mean test scores divided by the control group's test score standard deviation. For analyses with students from more than one grade, the mean control group standard deviation across grades was used to control for slightly larger variances in vertically equated test scores for multi-grade groups (see Appendix M).

^b Numbers in parentheses are the numbers of students in each group having valid test scores from the baseline 2006 administrations and the spring 2007 administrations.

The estimated *READ 180* impacts for all seven test scores are quite small—less than one standard/scale score unit. None are statistically significant ($p < 0.05$); and all effect sizes are quite small. There is no reason, therefore, to reject the hypothesis that the average achievement of the treatment and control groups was the same at the end of Year 1. In other words, *READ 180* participation did not have a significant impact on student achievement levels in reading or in the four core content areas at the end of the first year of the study.

Immediate Impact of *READ 180* at Each Grade in Year 1

The analyses of *READ 180* impact reported above were based on the total ITT sample of students in the sixth to eighth grades. The same analyses were also carried out separately for students in each of these three grades. Year 1 sixth-grade analyses results can be compared with results for sixth grade in Years 2–4 to see if there are any changes in the immediate impact of *READ 180* over time. The results for seventh and eighth-grade students in Year 1 allow a comparison of the impact of this targeted intervention across grades. The calculation of the immediate impact of *READ 180* for seventh and eighth-grade students in Years 2 through 4 is not possible due to their previous participation in *READ 180*.

Table 19 presents the results of the analyses of the immediate impact of *READ 180* on students in sixth grade.

Table 19: Immediate Year 1 Impact of *READ 180* on Spring 2007 Scores on Each Achievement Test—Grade 6

Test Score	Unadjusted Means		Adjusted Means		Est. Impact	Effect Size ^a	Signif. Level
	Control	<i>READ 180</i>	Control	<i>READ 180</i>			
ITBS Total Reading Standard Score	183.3 (287) ^b	184.9 (178)	183.8	184.3	0.5	0.04	0.665
ITBS Comprehension Stan- dard Score	179.7 (288)	180.0 (179)	180.3	186.9	-1.0	0.07	0.441
ITBS Vocabulary Standard Score	187.1 (290)	190.0 (180)	186.8	189.5	2.7	0.17	0.056
TCAP Reading/LA Scale Score	487.9 (372)	489.6 (228)	488.9	488.9	0.0	0.00	0.996
TCAP Mathematics Scale Score	494.4 (372)	494.9 (227)	492.3	494.7	2.4	0.08	0.279
TCAP Science Scale Score	185.9 (352)	187.9 (221)	186.6	187.6	1.0	0.07	0.423
TCAP Social Studies Scale Score	186.0 (349)	187.2 (222)	186.7	187.5	0.8	0.04	0.615

^a The method used to calculate effect size was Glass's Δ , the difference between treatment and control groups' adjusted mean test scores divided by the control group's test score standard deviation.

^b Numbers in parentheses are the numbers of students in each group having valid test scores from the baseline 2006 administrations and the spring 2007 administrations.

The estimated *READ 180* impacts in the sixth grade in Year 1 were not as small as they were for all grades. The impacts on the ITBS vocabulary subtest and TCAP mathematics were greater than two standard/scale score units. Again, however, none of the impacts are statistically significant ($p < 0.05$), although the impact on ITBS vocabulary approaches this level and the associated effect size, 0.17, is close to the small range of 0.2–0.5. Overall for sixth grade, there is no reason to reject the hypothesis that the average achievement of the treatment and control groups was the same at the end of Year 1. *READ 180* participation did not have a significant impact on student achievement levels in reading or in the four core content areas in the sixth grade in Year 1.

Similar analyses carried out for seventh and eighth-grade students yielded non-significant estimated impacts with effect sizes under 0.2. The complete results of the multi-level analyses of the Year 1 *READ 180* impact on these seven test scores for each grade can be found in Appendix B in Tables B8–B28.

Immediate Impact of *READ 180* at Sixth Grade in Year 2

The immediate impact of *READ 180* on ITBS and TCAP test scores for sixth-grade students in Year 2 was determined using the same multi-level model (see Appendix A) for the 693 sixth-grade students in the Year 2 ITT sample. The only differences were that the dependent variables

were the 2008 ITBS and TCAP test scores and the respective test score covariate was from the 2007 administrations. The results of these analyses are summarized in Table 20. The complete results of the multi-level analyses of the Year 2 *READ 180* impact on these seven test scores for sixth grade can be found in Appendix B in Tables B29–B35.

The estimated *READ 180* impacts in the sixth grade in Year 2 were of a similar size as they were in Year 1. There were more negative impacts favoring the control group. However, even the largest, for TCAP mathematics, was not statistically significant ($p < 0.05$). Overall for sixth grade, there is no reason to reject the hypothesis that the average achievement of the treatment and control groups was the same at the end of Year 2. *READ 180* participation did not have a significant impact on student achievement levels in reading or in the four core content areas in the sixth grade in Year 2. Thus, there was no variation in the impact of *READ 180* between Years 1 and 2; there were no impacts on any of the student outcome measures in either year.

Table 20: Immediate Year 2 Impact of *READ 180* on Spring 2008 Scores on Each Achievement Test–Grade 6

Test Score	Unadjusted Means		Adjusted Means		Est. Impact	Effect Size ^a	Signif. Level
	Control	<i>READ 180</i>	Control	<i>READ 180</i>			
ITBS Total Reading Standard Score	185.4 (295) ^b	182.9 (204)	184.4	183.7	-0.7	0.06	0.468
ITBS Comprehension Standard Score	182.3 (299)	179.4 (204)	181.8	180.0	-1.8	0.12	0.170
ITBS Vocabulary Standard Score	185.5 (299)	186.3 (210)	187.4	186.8	-0.6	0.04	0.639
TCAP Reading/LA Scale Score	497.1 (390)	495.6 (278)	494.7	496.5	1.9	0.06	0.407
TCAP Mathematics Scale Score	499.2 (390)	495.6 (278)	500.0	495.8	-4.2	0.15	0.070
TCAP Science Scale Score	187.0 (387)	185.9 (272)	186.5	186.3	-0.2	0.01	0.876
TCAP Social Studies Scale Score	186.7 (380)	182.8 (272)	185.8	183.5	-2.3	0.15	0.087

^a The method used to calculate effect size was Glass's Δ , the difference between treatment and control groups' adjusted mean test scores divided by the control group's test score standard deviation.

^b Numbers in parentheses are the numbers of students in each group having valid test scores from the baseline 2007 administrations and the spring 2008 administrations.

Long-Term (Two-Year) Impact of *READ 180* at Grades 7 and 8 in Year 2

The long-term impact of participating in *READ 180* for two years was examined in two different ways. First, a two-level model similar to the model described in Appendix A was used to estimate the difference between treatment and control students on spring 2008 ITBS and TCAP

test scores, controlling for their 2006 baseline scores, along with other student-level covariates. Since the students in these analyses were the 820 “stayers” in the seventh and eighth grades in Year 2, only one dummy variable indicating which students were in eighth grade was employed to control for within-school differences attributable to the student’s grade level. The school-level covariates remained the same. Table 21 presents the results of these cross-sectional analyses of the two-year impact of *READ 180*.

Table 21: Long-Term (Two-Year) Impact of *READ 180* on Spring 2008 Scores on Each Achievement Test

Test Score	Unadjusted Means		Adjusted Means		Est. Impact	Effect Size ^a	Signif. Level
	Control	<i>READ 180</i>	Control	<i>READ 180</i>			
ITBS Total Reading Standard Score	197.9 (398) ^b	198.4 (262)	198.0	198.1	0.1	0.01	0.967
ITBS Comprehension Standard Score	193.5 (405)	193.3 (260)	193.7	193.0	-0.7	0.03	0.639
ITBS Vocabulary Standard Score	201.8 (407)	203.3 (266)	201.5	203.3	1.8	0.08	0.280
TCAP Reading/LA Scale Score	505.8 (498)	507.5 (316)	505.3	506.7	1.4	0.05	0.446
TCAP Mathematics Scale Score	511.0 (498)	512.3 (315)	511.9	511.5	-0.4	0.01	0.871
TCAP Science Scale Score	184.0 (478)	185.0 (312)	184.4	184.7	0.3	0.02	0.782
TCAP Social Studies Scale Score	186.3 (474)	185.0 (310)	186.5	185.1	-1.4	0.11	0.129

^a The method used to calculate effect size was Glass’s Δ , the difference between treatment and control groups’ adjusted mean test scores divided by the control group’s test score standard deviation. For analyses with students from more than one grade the mean control group standard deviation across grades was used to control for slightly larger variances in vertically equated test scores for multi-grade groups (see Appendix M).

^b Numbers in parentheses are the numbers of students in each group having valid test scores from the baseline 2006 administrations and the spring 2008 administrations.

The estimated impacts of two years of participation in *READ 180* on the “stayers” in the seventh and eighth grades in Year 2 were not statistically significant ($p < 0.05$). There is no reason to reject the hypothesis that the average achievement of the treatment and control groups was the same at the end of Year 2. *READ 180* participation for two years did not have a significant impact on student achievement levels in reading or in the four core content areas in Year 2.

Similar analyses carried out separately for each of the two grades yielded non-significant estimated impacts with one exception. There was one significant difference in favor of the treatment group for ITBS vocabulary test scores. However, at a significance level of 0.05, this result may have also been due to chance since one would expect five percent of all hypothesis tests to

reject the null hypothesis of no difference when the null hypothesis is true. The complete results of the multi-level analyses of the *READ 180* impact on these seven test scores for the “stayers” in the seventh and eighth grades, together and separately by grade, can be found in Appendix B in Tables B36–B56.

The second method for examining long-term impacts of *READ 180* was to look for treatment and control differences in the average amount of growth in achievement from baseline to the end of Year 1 and from the end of Year 1 to the end of Year 2. This was accomplished by employing a three-level regression model. Level 1 represents a student’s test score as a function of the year in which it was measured, 2006, 2007, or 2008, or as the baseline score plus the growth in Year 1 plus the growth in Year 2. Level 2 represents the student’s baseline score as a function of membership in the *READ 180* treatment or control group and the same student covariates used in the previously described two-level analyses, and the growth in Years 1 and 2 as a function of whether the student was in the *READ 180* treatment or control group. Level 3 models the baseline score as a function of the same previously used school covariates. (See Appendix E for a complete specification of this model.)

Table 22 presents the key results of these longitudinal analyses of the two-year impact of *READ 180* on growth in achievement for the 820 “stayers” in the seventh and eighth grades in Year 2. The table presents the average growth in test scores for *READ 180* treatment and control students in Years 1 and 2 for each of the seven test scores. Underlined numbers indicate a significant ($p < 0.05$) difference between the amount of growth for treatment and control students. The bold number indicates the greater growth in each pair.

Table 22: Comparison of Growth in Mean Test Scores for *READ 180* and Control Students in Grades 7 and 8 in Year 2

Test Score	Year 1 Growth		Year 2 Growth	
	Control	READ 180	Control	READ 180
ITBS Total Reading Standard Score	<u>2.38</u>	<u>-0.02</u>	8.74	9.94
ITBS Comprehension Standard Score	<u>1.49</u>	<u>-1.79</u>	8.72	9.37
ITBS Vocabulary Standard Score	3.27	1.97	7.91	9.88
TCAP Reading/LA Scale Score	20.65	22.00	12.18	12.62
TCAP Mathematics Scale Score	15.32	12.33	11.72	12.45
TCAP Science Scale Score	5.01	3.90	-1.61	-0.63
TCAP Social Studies Scale Score	-0.16	1.52	<u>1.12</u>	<u>-1.98</u>

Note: Growth in mean test scores is based on the 319 students in the READ 180 treatment group and the 501 students in the control group from the Year 1 ITT sample who were enrolled in an MSRP school during the spring 2008 ITBS administration (the “stayers”). Significantly different ($p < 0.05$) growth in mean test scores in Year 1 or Year 2 between students in READ 180 and control groups are indicated by underlined numbers. The bold number indicates the greater growth in each pair.

In three of the 14 comparisons, treatment and control students demonstrated significantly different growth in achievement—for the ITBS Total Reading and Comprehension scores and for the TCAP Social Studies score. In each case the control students gained more than the students receiving *READ 180*. For ITBS Total Reading, control students gained over two standard score points in Year 1 while treatment students stayed the same. For ITBS Comprehension, control students gained one and a half points in Year 1 while treatment students lost over one and a half. For TCAP Social Studies, control students gained over a point in Year 2 and treatment students lost almost two.

Similar analyses carried out separately for each of the two grades yielded non-significant impacts on growth in average scores for seventh-grade students and several significant differences in eighth grade (see Table 23).

Table 23: Comparison of Growth in Mean Test Scores for *READ 180* and Control Students in Grade 8 in Year 2

Test Score	Year 1 Growth		Year 2 Growth	
	Control	<i>READ 180</i>	Control	<i>READ 180</i>
ITBS Total Reading Standard Score	<u>4.45</u>	<u>-0.46</u>	8.13	9.45
ITBS Comprehension Standard Score	<u>4.89</u>	<u>-0.35</u>	6.32	6.96
ITBS Vocabulary Standard Score	4.28	-0.21	8.70	10.94
TCAP Reading/LA Scale Score	24.60	24.32	15.96	17.69
TCAP Mathematics Scale Score	15.44	10.93	12.25	14.29
TCAP Science Scale Score	<u>7.01</u>	<u>2.22</u>	<u>-4.38</u>	<u>0.56</u>
TCAP Social Studies Scale Score	3.25	1.79	1.74	-1.28

Note: Growth in mean test scores are based on the 159 students in the READ 180 treatment group and the 241 students in the control group from the Year 1 ITT sample who were enrolled in an MSRP school during the spring 2008 ITBS administration (the “stayers”). Significantly different ($p < 0.05$) growth in mean test scores in Year 1 or Year 2 between students in READ 180 and control groups are indicated by underlined numbers. The bold number indicates the greater growth in each pair.

Significantly smaller growth was made by the treatment group in Year 1 for the ITBS Total Reading and Comprehension scores and for the TCAP Science score, while a significantly greater growth was made by the treatment group in Year 2 on the TCAP Science score. The complete results of the multi-level analyses of the *READ 180* impact on growth in these seven test scores for the “stayers” in the seventh and eighth grades, together and separately by grade, can be found in Appendix F in Tables F1 – F21.¹³

Differential Attrition in Immediate and Long-Term *READ 180* Impact Analyses

The numbers of students in the treatment and control groups in the analyses of immediate and long-term impact analyses are smaller than the numbers in the corresponding comparisons of baseline achievement levels. This is because not all of the students with valid baseline scores

¹³ If one compares the Year 1 growth estimates for grades 7 or 8 in Appendix F with Year 1 impact estimates in Appendix B, one will note that these estimates for any test score are not equal. This is not surprising since the students included in these two different sets of analyses are not the same. For example, Appendix B analyses are based on all students in the ITT group who remained in participating schools in Year 1, whereas the Appendix F analyses include only those students who remained in participating schools in Years 1 and 2 (the “stayers”). This difference also applies to the results in Appendices J and L.

also had valid scores from the spring 2007 and 2008 administrations at the end of Years 1 and 2. Although the level of attrition for both the treatment and control groups might or might not be similar, the important issue is whether the same type of students, especially in terms of their achievement levels, was lost from both groups. If not, one could argue that the estimated impacts were biased. That is, the treatment group may have lost students that would have scored higher (or lower) than the students lost from the control group.

This potential differential attrition was studied for the Year 1 ITT sample by comparing the average baseline 2006 test scores of the students who also had a spring 2007 score to the average of all students with baseline 2006 test scores, the difference being attributable to the attrition of students. This comparison was done for both the treatment and control groups.¹⁴ If this attrition effect were higher/lower in one group, this differential attrition would have to be acknowledged as possibly biasing the estimated impact of *READ 180* participation.

The results of the study of differential attrition for the estimated Year 1 immediate impacts may be found in Appendix C in Table C1.¹⁵ In summary, Table C1 shows that the effects of attrition in both treatment and control groups on baseline 2006 test scores did not exceed one standard or scale score point, and no differential treatment effects were statistically significant ($p < 0.05$). It seems reasonable to conclude that differential attrition was not a biasing factor affecting the interpretation of the estimated Year 1 immediate impacts of *READ 180*.

Differential Impacts of *READ 180* in MCLA and Non-MCLA Schools

The impact of MCLA on teacher and student outcomes is described later in this report. However, separate analyses were carried out with the *READ 180* ITT samples to determine whether the immediate and long-term impacts of *READ 180* in Years 1 and 2 were different in the MCLA treatment schools than in the MCLA control schools.

The dependent and independent variables and the covariates for addressing the research questions about the interaction of *READ 180* and MCLA are the same as those described in the above analyses of *READ 180* impacts, with one exception. An independent variable representing the participation of schools in the MCLA treatment was included in the analytical models for these analyses. This variable was included at the school level since schools were randomly assigned to the MCLA treatment or control condition. This model is specified in Appendix A.

The results of the analyses of the seven spring 2007 test scores for the *READ 180* ITT sample in Year 1 are presented in Table 24. The unadjusted and adjusted means are presented for the four combinations of *READ 180* treatment/control and MCLA treatment/control conditions. The estimated interaction effect is the difference between the estimated *READ 180* impact in MCLA treatment and control schools. A positive interaction effect means that the *READ 180* impact

¹⁴ The analytical method was a univariate ANOVA of baseline 2006 test scores, employing a 2x2 factorial design crossing the *READ 180* treatment/control condition with possession (yes/no) of a spring 2007 test score. The interaction of these two factors was tested for significance to determine whether or not there was a differential attrition effect.

¹⁵ Attrition effects are reported only for one of the ITBS test scores—Total Reading—since the other two subtest scores are very highly correlated with the Total Reading score, and results would be expected to be very similar.

was larger in MCLA control schools; a negative one means the *READ 180* impact was larger in MCLA treatment schools.

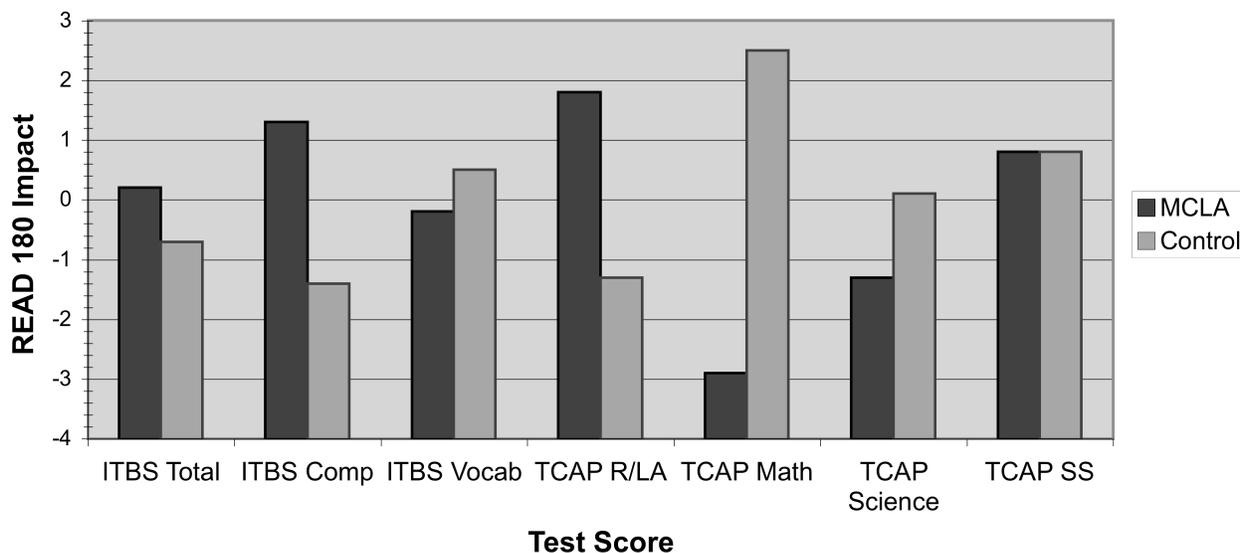
Table 24: Interaction of *READ 180* and MCLA Year 1 Impacts on Spring 2007 Scores on Each Achievement Test

Test Score	Unadjusted Means				Adjusted Means				Est. Interaction Effect ^a (A-C)-(B-D)	Effect Size	Signif. Level
	<i>READ 180</i>		Control		<i>READ 180</i>		Control				
	MCLA	Control	MCLA	Control	MCLA (A)	Control (B)	MCLA (C)	Control (D)			
ITBS Total Reading Standard Score	191.2 (231)	193.7 (280)	191.0 (371)	192.7 (341)	193.6	190.6	193.4	191.3	0.9	0.06	0.518
ITBS Comprehension Standard Score	187.8 (236)	187.5 (281)	185.9 (374)	187.5 (344)	189.7	184.3	188.4	185.7	2.7	0.14	0.168
ITBS Vocabulary Standard Score	196.1 (234)	200.2 (285)	196.3 (381)	197.8 (345)	197.4	197.7	197.6	197.2	-0.7	0.04	0.753
TCAP Reading/LA Scale Score	499.1 (311)	497.0 (353)	496.0 (512)	495.6 (460)	496.6	497.7	494.8	499.0	3.1	0.13	0.245
TCAP Mathematics Scale Score	496.1 (310)	506.8 (351)	498.4 (511)	501.7 (460)	495.9	504.4	498.8	501.9	-5.4	0.16	0.058
TCAP Science Scale Score	184.9 (305)	186.2 (338)	185.3 (503)	184.9 (412)	184.2	186.0	185.5	185.9	-1.4	0.09	0.388
TCAP Social Studies Scale Score	185.1 (306)	186.9 (338)	184.6 (495)	185.7 (411)	185.2	186.5	184.4	185.7	0.0	0.0	0.958

^a The formula for the estimated interaction effect reads, “The difference between the *READ 180* impact in MCLA schools and the *READ 180* impact in non-MCLA schools.” The method used to calculate effect size was Glass’s Δ , the difference between treatment and control groups’ adjusted mean test scores divided by the control group’s test score standard deviation (students in the *Read 180* control group and in an MCLA control school). For analyses with students from more than one grade the mean control group standard deviation across grades was used to control for slightly larger variances in vertically equated test scores for multi-grade groups (see Appendix M).

Figure 6 illustrates the magnitude and direction of these interactions. For example, on the ITBS Total Reading measure, the *READ 180* impact for the MCLA schools is 0.9 points higher (0.2– (-0.7)) in the MCLA schools than it is in the non-MCLA or control schools.

Figure 6: *READ 180* Year 1 Impact in MCLA Treatment and Control Schools



These interaction effects range between -5.4 and 3.1, an impact that was slightly more than five-scale score points larger on spring 2007 TCAP mathematics scores in the MCLA control schools and three-scale score points larger on TCAP reading/language arts scores in MCLA treatment schools. Clearly these results are mixed, although the *READ 180* impact is larger in MCLA treatment schools on reading measures (except vocabulary) and larger in MCLA control schools in the other content areas. However, none of the interaction effects are statistically significant ($p < 0.05$), although the effect for the TCAP mathematics scores approached this level. Also, all of the effect sizes are less than 0.2. In summary, it seems reasonable to conclude that the *READ 180* impact on student achievement did not vary significantly between MCLA treatment and control schools in Year 1.

The interaction of the *READ 180* and MCLA impacts was also analyzed for each grade separately. The results for the sixth grade are presented in Table 25. Again, the results were mixed and non-significant for the most part. The *READ 180* impact on the spring 2007 TCAP science scores was significantly larger in MCLA control schools than in MCLA treatment schools ($p < 0.05$), and the effect size was 0.32. The larger *READ 180* impact on ITBS vocabulary scores in MCLA treatment schools was not statistically significant, but the effect size was 0.21. It is interesting to note that the pattern of (non-significant) larger *READ 180* impacts in the MCLA treatment schools for reading measures and larger impacts in the control schools for non-reading measures was also present in the sixth-grade results.

Table 25: Interaction of Year 1 *READ 180* and MCLA Impacts on Spring 2007 Scores on Each Achievement Test—Grade 6

Test Score	Unadjusted Means				Adjusted Means				Est. Interaction Effect ^a (A-C)-(B-D)	Effect Size	Signif. Level
	<i>READ 180</i>		Control		<i>READ 180</i>		Control				
	MCLA	Control	MCLA	Control	MCLA (A)	Control (B)	MCLA (C)	Control (D)			
ITBS Tot Reading	185.7 (79)	184.3 (99)	183.4 (145)	183.3 (142)	185.0	183.7	183.8	183.8	1.3	0.11	0.529
ITBS Comprehension	181.4 (80)	178.9 (99)	181.2 (146)	178.2 (142)	180.4	178.4	181.2	179.4	0.2	0.01	0.947
ITBS Vocabulary	190.3 (80)	189.7 (100)	185.6 (146)	188.6 (144)	191.2	188.0	186.7	186.7	3.2	0.21	0.258
TCAP Reading	491.3 (110)	488.0 (118)	489.4 (189)	486.4 (183)	488.4	489.5	487.8	490.2	1.3	0.05	0.781
TCAP Mathematics	493.6 (110)	496.1 (117)	497.2 (189)	491.4 (183)	493.6	495.4	494.3	490.1	-6.0	0.19	0.166
TCAP Science	186.2 (108)	189.5 (113)	187.0 (187)	184.8 (165)	188.3	186.9	189.9	183.1	-5.4	0.32	0.037
TCAP Social Studies	186.9 (109)	187.6 (113)	187.4 (184)	184.5 (165)	186.7	187.6	187.6	185.0	-3.5	0.19	0.240

^a The formula for the estimated interaction effect reads, “The difference between the *READ 180* impact in MCLA schools and the *READ 180* impact in non-MCLA schools.”

Separate analyses conducted for the seventh and eighth grades also yielded only a few interactions worth noting, but they did all favor the *READ 180* impact in MCLA treatment schools. Two statistically significant ($p < 0.05$) interactions were found that supported a larger *READ 180* impact in MCLA treatment schools on ITBS comprehension scores for eighth-grade students (effect size = 0.33) and on TCAP reading/LA scores for seventh-grade students (effect size = 0.40). Another two interactions had effect sizes slightly above 0.20, but were not statistically significant ($p < 0.05$). These two suggested a larger *READ 180* impact in MCLA treatment schools on TCAP reading/LA scores for eighth-grade students and on TCAP science scores for seventh-grade students. The complete results of the multi-level analyses of the Year 1 *READ 180* impact for MCLA treatment and control schools for students in all three grades and each grade separately can be found in Appendix D in Tables D1–D28.

Additional analyses of the interaction between *READ 180* and MCLA were carried out on student outcome measures obtained at the end of Year 2. One set of seven analyses examined the interaction of the immediate impact of *READ 180* and MCLA for the Year 2 sixth-grade ITT sample, looking at the spring 2008 ITBS and TCAP scores and controlling for the fall 2007 ITBS and spring 2007 TCAP scores. A second set examined the interaction of the long-term *READ 180* impact and MCLA for the Year 2 seventh and eighth-grade “stayers,” looking at the 2008 Spring ITBS and TCAP scores controlling for the fall 2006 ITBS and spring 2006 TCAP scores. None of the interaction effects in these 14 analyses were statistically significant ($p < 0.05$). The complete results of these additional multi-level analyses can be found in Appendix D in Tables D29–D42.

There do not appear to be any clearly interpretable patterns in the differences in *READ 180* impact for MCLA treatment and control schools. In Year 1 there was a suggestion of larger *READ 180* impacts on reading measures in MCLA treatment schools versus larger impacts on non-reading measures in control schools in the sixth and seventh grade analyses as well as the analyses based on all grades. However, of the 42 interaction effects tested for Years 1 and 2, only three (7%) were statistically significant. Using a significance criterion of $p < 0.05$, five percent (or 2) of the 42 tests would be expected to be found significant by chance.

Conclusions

The lack of any significant immediate impacts of *READ 180* participation in Year 1 continued into Year 2. There were no immediate impacts in Year 2 on sixth-grade students and only one out of 21 long-term impacts was significant for the seventh and eighth grades based on cross-sectional analyses of spring 2008 test scores. Long-term impacts based on a growth model revealed seven significant growth coefficients, six of which were in favor of the control group. Finally, an examination of the interaction between the impacts of *READ 180* and the whole-school intervention, MCLA, yielded no clearly interpretable patterns in the differences in *READ 180* impact for MCLA treatment and control schools.

Further Analyses

In an earlier footnote, it was noted that three-level regression models were explored to include the characteristics of teachers linked to students for each core content area. These analyses

were sufficiently complex and the results sufficiently equivocal to lead to the decision to omit their results from any further reporting. In addition, they did not result in different estimates of *READ 180* impact. As a consequence, it was also decided to omit analyses designed to determine if *READ 180* impacts are moderated by teacher characteristics. (Moderation by school characteristics was not investigated in the two-level models due to the very low levels of between-school variation.) Exploratory analyses are still planned that will study the relationship between *READ 180* impact and the amount of *READ 180* instruction received by students randomly assigned to the treatment group.

V. Evaluation of the Implementation of the Whole-School Intervention: Years 1 and 2

The implementation evaluation of the MCLA whole-school intervention in Years 1 and 2 addressed three overarching research questions:

1. To what degree did the implemented MCLA treatment match the intended program standards and features?
2. What contextual district and school-level factors influenced the implementation of the MCLA program?
3. How did the professional development events, materials, or structures present in the control schools compare to what was present in the treatment schools?

Specific research questions about the implementation of the two-year whole-school intervention include:

- What was the Year 2 MCLA classroom instructional model?
- What types and amount of professional development were provided to teachers, principals, literacy coaches, and MCLA instructors?
- What proportion of teachers received and participated at different levels of professional development (e.g., how many used program materials or completed the MCLA course?)
- What types of coaching support was provided to teachers?

The development of the IC Map (Cooter et al, 2008) was completed in Year 2 and will be used in Years 3 and 4 of the MSRP to measure teachers' implementation of MCLA strategies in the classroom. In Year 2, RBS measured classroom implementation through a teacher survey, focus group interviews in the fall and spring, and an analysis of logs maintained by the literacy coaches.¹⁶

RBS reviewed information contained in course syllabi, MCLA instructor templates, focus group interview transcripts, and CRC inventory lists to address the research question about the extent to which the MCLA treatment matched its design. Next, RBS conducted interviews with

¹⁶ Rui and Feldman (2008) conducted a validity study of the MSRP classroom observation protocol developed by Feldman and Feighan (2007) and will present results at the annual meeting of the 2009 American Education Research Association.

the eight participating MSRP principals and collected survey data from 169 teachers to provide insight into factors influencing implementation and for information about any related professional development events offered in the control schools. Evaluators also examined MCLA course attendance sign-in sheets, coaching logs, CRC logs, and information collected through MCLA course observations for answers about the types of professional development provided and level of program participation. Finally, RBS analyzed coaching logs, focus group interviews, and responses on two additional teacher surveys for further detailed information about the type of coaching services that were provided to participants. Figure 7 summarizes the relevant sources of data used in assessing the Year 2 implementation of MCLA.

Figure 7: Data Sources Linked to Research Questions—MCLA, Year 2

Research questions	Measures/Data Sources									
	Surveys/Logs				Classroom Observations			Record Review		
	Teacher	Coach	District	Developer	Evaluator	Developer/coach	Other district staff	PD attendance records	Lesson plans	Student attendance
What was the level of implementation and variability of MCLA professional development and support for teachers, coaches, and building principals in Year 2?										
<i>Professional development for teachers</i>										
Type/amount of PD provided to teachers	X	X			X	X		X		X
Proportion of teachers at different levels of PD	X							X		
Proportion of teachers at adequate level of PD	X				X	X		X		
Types/amount of coaching provided to teachers	X	X				X				
Proportion of teachers at different levels of coaching	X	X								
Proportion of teachers at adequate level of coaching	X	X								
<i>Professional development for coaches/other relevant staff</i>										
Type/amount of PD provided to coaches		X	X					X		
Proportion of coaches at different levels of PD		X								
Type/amount of PD provided to school principals				X				X		
Proportion of school principals at different levels of PD				X				X		
Type/amount of PD provided to district leaders			n/a					n/a		
Proportion of district leaders at different levels of PD			n/a					n/a		

Research questions	Measures/Data Sources									
	Surveys/Logs				Classroom Observations			Record Review		
	Teacher	Coach	District	Developer	Evaluator	Developer/coach	Other district staff	PD attendance records	Lesson plans	Student attendance
What was the level of implementation and variability of classroom instruction in Year 2?										
Proportion of teachers with access to materials and resources	X	X	X							
Proportion of teachers who implemented literacy strategies (CAPs)	X	X	X							
Proportion of teachers who implemented the model at adequate level	X		X							
What did the counterfactual look like in Year 2?										
Proportion of teachers at control schools reporting literacy-related PD at follow-up	X									

Contextual Factors in Control and Experimental MSRP schools

In May 2008, with assistance from MCS, RBS administered a survey to content area teachers working in the eight MSRP schools. Teachers were asked to think about the 2007–2008 school year when answering questions about:

- How many hours of professional development in specific topic areas they had received
- How prepared they felt to engage in a set of 24 specific literacy activities
- How often they had implemented those literacy strategies

A total of 169 teachers completed the survey: 101 (59.8%) respondents worked in control schools and 68 (40.2%) respondents were from experimental schools. Among the 68 experimental teachers, 47 (69.1%) had participated in MCLA and 21 (30.9%) had not. Only one difference was found in teachers' reported participation in various professional development topic areas: MCLA teachers were more likely than control teachers or non-MCLA teachers in experimental schools to report having had training in the area of literacy integration during the 2007–2008 school year ($F = 18.5, df = 2, 164; p < .05$). As Table 26 shows, 41.2 percent of teachers in the four MCLA schools indicated receiving more than 32 hours of professional development during the year in the area of literacy integration, compared with 15.8 percent of control teachers.

Table 26: Professional Development Participation in Literacy Integration in the Past Year among Control and Experimental Teachers, May 2008

Participated in Professional Development in the Area of Integrating Literacy in the Classroom	None	1 to 8 Hours	9 to 32 Hours	32+ hours
Control group (N=101)	6.9	51.5	23.8	15.8
Experimental group (N=68)	0.0	20.6	36.8	41.2

The survey analysis also found that over one-third (36.8%) of experimental teachers and 29.7 percent of control teachers had received no professional development in the past year to ad-

dress the needs of ELL students/students from diverse backgrounds. Finally, results also showed that MCLA teachers reported using graphic organizers more frequently than non-MCLA teachers at experimental schools or control teachers ($F = 3.89$, $df = 2$, 162; $p < .05$) and that MCLA teachers were more likely than others to report feeling prepared to use 10 of 24 literacy strategies identified on the survey ($F = 5.92$, $df = 2$, 158; $p < .05$; see Appendix N-2A for the full summary of teachers' responses and Appendix N-2B for the survey instrument).

In fall 2007, evaluators also conducted observations of 48 classrooms in the eight MSRP schools to document contextual factors that might affect implementation as well as the extent to which control and experimental teachers implemented literacy strategies in their content classes. A team of 12 researchers observed 22 control school classrooms and 26 treatment school classrooms. Nineteen of the treatment teachers (73%) were MCLA participants, while the other seven teachers (27%) in the treatment schools had not enrolled in the program. The grade levels represented were relatively even as were the distribution of content classes. See Appendix N-3 for a full report and Appendix N-4 for the classroom observation protocol. Overall findings include the following:

- Observers recorded a greater availability of books in MCLA treatment classrooms than in control classrooms ($F = 11.75$, $df = 1$, 46; $p < .05$).
- Observers rated the climate of respect for students' experiences and ideas more highly in the MCLA treatment classrooms than in control classrooms ($F = 7.86$, $df = 1$, 45, $p < .05$). No other differences were noted in the classrooms' physical environment or social climate.
- Similar to Year 1, there were no differences in student engagement level between treatment schools and control schools in the fall 2007 observations. Overall, students in the observed classes had relatively high levels of participation/engagement.
- Observers noted the use of at least one literacy strategy in 73.1 percent ($N = 19$) of the MCLA teachers' classes compared with 54.5 percent of control classes ($N = 12$). Ten MCLA treatment teachers used three or more literacy strategies, compared with four control teachers, and MCLA participants tended to use more literacy strategies than other teachers, although sample sizes across the three groups of teachers (MCLA, non-MCLA in treatment schools, and control) are insufficient to test for statistical significance.

Professional Development Model for Teachers as Implemented

The following section explores the extent to which the implementation of the teacher and principal courses in Years 1 and 2 approximated the intended MCLA design. Next, the report provides details about participation and retention rates in the teacher and principal courses and then summarizes participants' use of the CRC before reporting on teachers' collaboration with literacy coaches. RBS next presents the results from classroom observations for insight into the classroom level of MCLA implementation. Finally, the analysis moves to a summary of MCLA implementation that includes ratings for each of the four participating schools in Years 1 and 2.

MCLA Course Content

Figure 8 summarizes the topics addressed in the two-year MCLA course cycle as indicated on the syllabi for each of the four semesters (see Appendix N-5 for the Year 2 course syllabi). The two-year course initially exposed teachers to specific literacy strategies aimed at building vocabulary, fluency, and comprehension. The final semester reviewed the previously learned strategies and emphasized when during instruction (before, during, or after) they were most appropriately implemented.

In Years 1 and 2, teachers were responsible for completing a total of 14 CAPs, two fewer than developers initially planned. During the fall semester of Year 2, CAPs focused on previewing text, concept maps, word walls, and comprehension monitoring skills, while the spring semester served as a review. Figure 8 summarizes the CAP assignments for Years 1 and 2.

Figure 8: MCLA Years 1 and 2 Course Topics

Fall 2006	Spring 2006	Fall 2007	Spring 2008
Gradual Release of Responsibility: Vygotsky's Zone of Proximal Development Preteaching vocabulary in Core Subjects Introduction to Word Maps Comprehension and Vocabulary Learning: Student Generated Questions Comprehension: Thinking Maps and Semantic Features Analysis Comprehension: Adapting Instruction for Special Needs Students Modeling Fluency: Read Alouds Fluency: Choral Reading Fluency: Reading in Pairs and Alternative Activities for Special Needs Learners Cooperative Learning: Think-Pair-Share, Jigsaw, and Working in Dyads	Improving Vocabulary Knowledge, Writing & Comprehension Comprehension Workshop: Oral Retellings Getting to know the National Geographic "Leveled" Books at Your School Improving Vocabulary Knowledge, Writing & Comprehension: Oral Retellings Retelling Using a Thinking Map/ Graphic Organizer Special Session at Sherwood: Overview of <i>National Geographic</i> Materials Written Retellings (paragraphs) The Memphis Pyramid: Review of Three-Level Retelling	Previewing Text with emphasis on student generated questions "THIEVES" Previewing Text Strategy Reciprocal Teaching Explicit Vocabulary Instruction: Concept Maps Categorizing Using Group-generated Content Word Walls Repeated Exposures to "Marinate" Students in New Content Vocabulary Comprehension Monitoring: REQUEST Technique Question-and-Answer Relationships Revisited	What Teachers Can Do to Improve Students' Vocabulary Knowledge: Before, During, and After Strategies What Teachers Can Do to Improve Students' Comprehension: Before, During, and After Strategies What Teachers Can Do to Improve Students' Reading Fluency: Before, During, and After Strategies

Data source: MCLA syllabi

MCLA developers and their team of content-specialist writers created a template for instructors to follow during each session. Directing instructors on how to conduct an activity with participating teachers, the templates included a rubric to guide the activity, a scoring sheet, an out-of-class assignment, and the CAP. See Appendix N-6 for an example of a fall 2007 instructor

template on the topic of previewing text to improve comprehension. These templates were constructed during Years 1 and 2 of the project, and they will continue to be fine-tuned in Years 3 and 4 with teachers from the second cohort of MCLA schools.

Four of the original five Year 1 instructors taught the four content area classes in both the fall and spring semesters of Year 2. Results from interviews conducted with instructors in May 2008 reveal that while course attendance was generally high, instructors found that the level of motivation and enthusiasm waned among teachers who had been in the program for the full two years. On the other hand, instructors stated that class presentations went particularly well during Year 2, and three of the four instructors agreed that MCLA activities were explicitly linked to state and district standards. Although instructors relied on their own system or philosophy in assigning MCLA grades to participants, each made only slight/occasional modifications to the instructor template/scripts and cited a similar structure and flow of activities across the different content area classes.

Figure 9: CAP Assignments, Years 1 and 2

Year 1	Year 2
<p><i>Fall 2006</i></p> <ul style="list-style-type: none"> • Vocabulary knowledge • Student-generated questions • Read alouds • Think-pair-share • Semantic features <p><i>Spring 2007</i></p> <ul style="list-style-type: none"> • Using the <i>National Geographic</i> leveled books at your school • Oral retelling • Written Retellings of Expository Text 	<p><i>Fall 2007</i></p> <ul style="list-style-type: none"> • T.H.I.E.V.E.S. (previewing text and generating questions) • Academic word walls • Question-answer-relationship (QAR) <p><i>Spring 2008</i></p> <ul style="list-style-type: none"> • Vocabulary Instruction: Before, During, and After Reading • Reading Fluency: Before, During, and After reading • Comprehension Monitoring: Before, During, and After Reading

Data source: MCLA syllabi

Instructor templates across the four content areas followed the same general pattern in both the fall and spring, although activities, articles, and presentations were content-specific. For example, fall scripts contained identical “non-negotiable core” topics across all content areas (e.g., small-group activities), but also included at least an hour of time devoted to applying literacy strategies to a particular content area. The templates integrated components of a pedagogical model developed by the Center for Research in Education, Diversity and Excellence (CREDE) at the University of California, Berkeley, that emphasized small-group “joint-productive activities” (JPAs) rather than whole-group instruction or lecture, and time for reflection.

Twelve MCLA course observations were conducted during nine (47%) of 19 weeks that the Year 2 course was offered to participating teachers. Specifically, evaluators observed two sessions in mathematics, three in ELA, three science sessions, and four social studies sessions spread across Year 2. Observers found that JPAs were used consistently across the content areas and that teachers recorded reflections in 10 of 12 observed sessions. Most observed classes employed a mix of instruction and practice in use of literacy strategies. Instructors often introduced

new concepts through articles and handouts read silently or aloud in class, and five of the observed sessions included strategy modeling, guided practice in the strategy, and then independent practice. (See Appendix N-7 for a more detailed summary of these MCLA course observations.)

The fall semester was organized around three themes: previewing text with an emphasis on student-generated questions, explicit instruction in vocabulary, and comprehension monitoring. The spring semester, serving as a review of strategies taught in the prior three semesters, was organized into dimensions that embodied the main purpose of MCLA: improvement of vocabulary, comprehension, and fluency. Each dimension was addressed in a three-week cycle: the key strategies were reviewed during the first week, teachers were to model the strategy through presentation lessons during the second week, and then teachers were to work collaboratively to develop their CAPs during the third week. Participants did not learn new strategies during the second semester, but rather they practiced the classroom applications of previously learned strategies. Observers found that MCLA material presented was linked explicitly to district standards or student performance indicators in only one of the 12 observed sessions. Research supporting the use of each strategy was occasionally presented in class; however, this research was generally presented in the form of a handout or article and was rarely addressed by the instructor during class.

Finally, the use of the CRC was also encouraged in the MCLA course through promotion on the syllabi and discussions about how to integrate CRC materials into different lessons. Participants were urged to visit the MCLA website during three class sessions and were directed to websites with additional material (such as graphic organizers and state lists of academic vocabulary words) during other sessions. A content-area teacher introduced the Visual Thesaurus, a software program purchased by MCS, during two class sessions; one of these discussions involved an active demonstration of the program.

The following narrative provides results about the proportion of individuals that participated in different levels of professional development. Specifically, it describes attendance in the MCLA teacher course and principal fellowship, use of the CRC, and extent to which teachers worked with a literacy coach.

MCLA Course Participation

In Year 2, MCLA professional development was expanded to include special education resource teachers in addition to teachers working in the core content areas. The program continued to be offered to teachers working in Year 1 MSRP schools and began officially on August 9, 2007, at the annual kickoff event at FedEx's state-of-the-art arena in downtown Memphis. Seventy-four (86%) of the 86 enrolled MCLA teachers completed the fall semester course and 66 (89%) of the 74 participants finished the spring semester course and its requirements. (Thirteen individuals dropped the class during the winter break while five new teachers joined). In all, 46 teachers completed the 2-year MCLA cycle.

Instructors teaching the MCLA course in Year 2 provided 56.5 hours of professional development to participants. Program staff provided a daylong introductory session, nine fall semes-

ter evening course sessions, nine spring semester classes,¹⁷ and a final daylong commencement session that marked the conclusion of the whole-school intervention. Evening sessions typically ran from 4:15 to 6:30 p.m. and the kickoff and laureate ceremonies lasted in duration from approximately 9 a.m. to 3 p.m. Approximately 85 hours of classes were provided in Year 1, so a total of approximately 141.5 hours of professional development class time was provided to teachers during the two-year cycle. This figure is almost 40 hours shy of the target goal to provide 180 hours of MCLA training; however, this figure excludes time spent working with literacy coaches, which was considered professional development time in Year 2.

In October 2007, district data showed that there were 115 full-time teachers in the four schools whose primary content area was ELA, reading, mathematics, science, social studies, or special education. Of these teachers, 48 taught ELA or reading, 23 taught mathematics, 21 taught science, 18 taught social studies, and five taught special education/CDC. Special education teachers enrolled in one of the four MCLA core content classes offered and were tracked for this analysis according to their MCLA course attendance and not status as a special education teacher. The analysis of MCLA participation shows that 74 of 115 (64.3%) teachers eligible for MCLA attended the course and received a stipend for participation, which represents an increase from 51 percent of eligible teachers in Year 1. As Table 27 shows, participation rates were highest in Year 2 among mathematics teachers, where 78.3 percent of eligible mathematics teachers in the four schools did participate in MCLA. The proportion of eligible ELA/reading teachers participating in the program was lowest at 58 percent, although actual numbers of participating ELA/reading teachers were highest compared with the numbers of teachers in other content areas. (During Year 1, the proportion of teachers participating by content area was highest among those teaching ELA).

Table 27: Number of MCLA Course Completers by Content Area, Fall 2007

Content Area	Number of Full-time Teachers Eligible to Participate	MCLA Completers	
		#	%
ELA/READ 180	48	25	58.3
Mathematics	23	21	78.3
Science	21	12	57.1
Social studies	18	16	61.1
Total	110*	74	

* Number excludes five special education teachers.

RBS determined MCLA eligibility based upon the cross-validation of the number of core content and special education teachers working in the four participating schools in October 2007 according to the information provided by MCS and collected by RBS. The number and percentage of participating teachers by school, shown in Table 28, indicate that participation was highest in Riverview and Sherwood Middle Schools, where nearly three-quarters of eligible teachers participated in MCLA, and lowest at A. Maceo Walker (48.4%) and Hamilton (62.5%). These fig-

¹⁷ Two science sessions and one ELA session were canceled during the spring semester session due to inclement weather.

ures are similar to those reported in Year 1, during which MCLA participation was highest at Riverview and lowest at A. Maceo Walker.

Course completion rates were moderate to high: 74 (86%) of the 86 originally enrolled teachers completed the fall semester course. Thirteen (17.6%) fall completers withdrew from the MCLA course during the winter break, while five new teachers enrolled during that time. A total of 66 teachers finished the spring course; 61 (70.9%) of whom had also completed the fall course.

In May 2008, evaluators attempted to interview 17 teachers known at the time to have withdrawn from MCLA and learned that six individuals had retired, relocated, took maternity leave or experienced scheduling conflicts such as afterschool tutoring or Teach for America obligations. Evaluators interviewed seven (63.6%) of the 11 remaining respondents, and found that three had employment or graduate school-related scheduling conflicts, one suffered from medical problems, and three stated that they were either busy or overwhelmed with other work commitments, such as after-school coaching or balancing their workload.

Table 28: Number and Percentage of MCLA Participants by School, Year 2

School	Fall 2007 (N = 74)				Spring 2008 (N = 66)	
	Number of Participants	Percentage of all MCLA Participants	Number of Eligible Content Teachers in School*	Percent of Eligible Teachers in MCLA	Number of Participants	Percentage of all MCLA Participants
A. Maceo	16	21.6	33	48.4	13	19.7
Hamilton	15	20.3	24	62.5	13	19.7
Riverview	14	18.9	19	73.7	14	21.2
Sherwood	29	39.2	39	74.4	26	39.4
Total	74	100%	115		66	100%

Table 29 summarizes teachers' MCLA participation by content area course for Years 1 and 2. Since some teachers attended a different weekly content session than the one for which they officially registered, RBS totals on participant attendance by content area differ slightly from an MCS list of completers, which is based upon a teacher's registered content area (and not necessarily the class they attended routinely). The total number of participants in the RBS and MCS datasets, however, is identical.

Attendance at the course sessions was high: Virtually all MCLA participants (97.2%) attended seven or more of the ten sessions offered in fall 2007. Specifically, 59 (79.7%) of the 74 teachers who completed the fall course attended eight or more of the ten total sessions, 13 (17.6%) attended seven sessions, and two teachers attended five and six sessions (2.7%). None of the completers attended fewer than five sessions. Nearly a third (32.2%) of the participants attending 80 percent or more sessions had attended every session; in fact, the rate of perfect attendance would have likely been higher if not for a mandatory special education in-service training scheduled the same day as the MCLA kickoff. Attendance was lower in the spring but still moderate overall: 41 (62.1%) of the 66 spring completers attended eight or more of the ten ses-

sions. Twelve (18.1%) attended six or seven times and five (7.5%) attended five or fewer sessions. Of the 41 high attenders, eight (12.1%) teachers had perfect attendance.

Table 29: Number of MCLA Course Participants by Content Area, Year 2

Content area	Number of Participants			
	Year 1		Year 2	
	Fall 2006 (N=69)	Spring 2007 (N=72)	Fall 2006 (N=74)	Spring 2007 (N=66)
ELA/READ 180	28	29	25	21
Mathematics	18	18	21	20
Science	12	15	12	10
Social studies	11	10	16	15

Data source: MCLA course attendance sheets provided by instructors.

Table 30 summarizes the percentage of participants by content area with high MCLA attendance, defined as having attended 80 percent or more of the total number of sessions offered in the fall and spring. (Due to weather cancellations in the spring semester, there were eight classes offered in science, nine classes in ELA, and 10 classes in the mathematics and social studies content areas). As the table shows, the percentage of teachers attending the science content area sessions decreased from 91.7 percent in the fall to 60 percent in the spring even though two spring sessions were cancelled: six of ten science teachers attended seven or eight sessions, while four teachers attended six or fewer of the eight spring sessions offered in that content area. Mathematics teachers had the greatest proportion of high attenders in the spring (90%).

Table 30: Percentage of MCLA Teachers Attending 80 Percent or More Sessions by Content Area

	Fall 2007			Spring 2008			Year 2	
	Total Completing Content Classes	High Attendance		Total Completing Content Classes	High Attendance		Overall High Attendance in Year 2*	
		#	%		#	%	#	%
ELA/READ 180	25	18	72.0%	21	15	71.4%	12	57.1
Mathematics	21	16	76.2%	20	18	90.0%	15	83.3
Science	12	11	91.7%	10	6	60.0%	7	63.6
Social Studies	16	14	87.5%	15	12	80.0%	7	63.6
Total	74	59		66	51		40	65.6

Data source: MCLA course attendance sheets provided by instructors.

* A total of 61 participants completed both fall and spring semesters: 21 ELA/reading teachers, 18 mathematics teachers, 11 science teachers, and 11 social studies teachers.

MCLA Principal Fellowship Course Participation

MCLA developers invited building principals and other administrative staff to participate in the graduate level course, “Advanced Reading Instruction for the Special Learner: MCLA Principals’ Fellowship” (see Appendix N-8 for the course syllabus). The class met for six sessions in the fall and four sessions in the spring. Attendance data provided by the developer indicate that all seven participants achieved perfect fall semester attendance. Participants were four building principals, an instructional facilitator, and two assistant principals from the same school. The same participants achieved perfect attendance in the spring (two principals met with the MCLA developer individually after the fellowship meeting to make up for a missed class). Attendance and participation in the principal fellowship increased dramatically between Year 1 and 2, after the developer changed the frequency of meetings from weekly to monthly sessions and included the assistant principals and other building administrators.

Formal interviews conducted with the principals in May 2008 corroborate their perfect attendance and reveal a positive experience overall. Respondents stated that the fellowship afforded them an opportunity to discuss the issues related to carrying out the whole-school intervention and allowed them to learn from each other. In addition, the course provided staff development tools that the principals could use with teachers, research about current literacy trends, and information about using data to track whether instructional efforts are improving student test scores.

Curriculum Resource Center (CRC) Use

In addition to funding the professional development course and literacy coaching services, the MSRP grant provided participants at the four MCLA schools with a CRC that housed an array of reading materials and themed resources for use with their students. Inventory records include the following types of materials: *National Geographic* leveled text thematic sets, the *Math Matters* series by Grolier, the *TIME* Secondary Science series, and Hampton Brown’s *Picture It* among others. Although the CRC was fully operational in each school at the start of Year 2, an analysis of checkout logs reveals a steep decline CRC use from Year 1 when a majority of MCLA participants had used the materials at least once.

Table 31 summarizes the number of CRC items checked out in Year 2, the number of individual MCLA teachers checking out those items, and use among MCLA participants in spring 2008. In Year 2, exactly half (N=33) of the sixty-six MCLA teachers checked out at least one resource, compared with 59 of 70 (84.3%) teachers in Year 1. The 33 teachers borrowed a total of 127 separate items in Year 2, compared with 235 items borrowed during the previous year. One explanation for the decline in checkout figures is that only three of the four schools submitted CRC data for Year 2; however, interviews conducted with each literacy coach confirm that CRC use was lower compared with Year 1 levels.

Year 2 results show that between 3 and 16 individual teachers in the four schools borrowed CRC materials, representing half of all MCLA participants. All (100%) participants from A. Maceo Walker used the CRC at least once compared with less than one-quarter (21.4%) of teachers at Riverview. It is important to note that a “resource” may include a set of materials rather than

an individual educational item. For example, RBS counted a teacher’s use of six *National Geographic Money and Time* books and the related disc and transparency as one item.

Table 31: CRC Resource Usage by School, Year 2

School	Number of MCLA Teachers in Spring 2007 (N = 66)	Number (Percentage) of MCLA Teachers Checking Out Resources (N =33)	Number of Resources Checked Out (N=127)
A. Maceo Walker	13	13 (100%)	91
Hamilton	13	*	*
Riverview	14	3 (21.4%)	5
Sherwood	26	16 (61.5%)	31

Data source: CRC checkout logs, Year 2

* Logs from Hamilton Middle School were not submitted

Teachers continued to primarily use the *National Geographic* materials in Year 2 as they had in Year 1; however, some teachers also used new materials provided by the grant team such as *Building Fluency Reader’s Theater* by Teacher Created Materials or resources lent temporarily by a literacy coach from a privately owned collection (i.e., trade books). Table 32 summarizes the type of resources teachers used in the CRC during Years 1 and 2.

Table 32: CRC Resource Usage by Category, Years 1 and 2

Category	Year 1 (N=235 items)		Year 2 (N=127 items)	
	Number	Percentage	Number	Percentage
National Geographic– Life Science/Human Body	45	19.1	11	8.7
Social Studies– Various Materials	25	10.6	3	2.4
National Geographic– U.S. History and Life	30	12.8	5	3.9
National Geographic– Earth Science	19	8.1	20	15.7
National Geographic– Life Science	17	7.2	0	0.0
Science– Various Materials	15	6.3	2	1.5
National Geographic– Math	21	8.9	6	4.7
Professional Library	13	5.5	0	.00
National Geographic– Science Theme Sets	13	5.5	0	0.0
Mathematics– Various Materials	10	4.2	8	6.3
National Geographic– Social Studies Theme Sets	6	2.5	0	0.0
National Geographic– Ancient Civilizations	4	1.7	4	3.1
National Geographic– Physical Science	4	1.7	22	17.3
Professional Development	3	1.3	0	0.0
Science Matters/Visual Science Encyclopedia	3	1.3	0	0.0
Science Theme Sets	5	2.1	0	0.0
U.S. Regions	2	0.8	1	0.7
Teacher Created Materials	n/a	n/a	9	7.1
ELA– trade books, Janet Allen themed sets	n/a	n/a	23	18.1
Literacy Coach Materials (e.g., books, study guide)	n/a	n/a	8	6.3
Unknown type	n/a	n/a	5	3.9
Total	235	100%	127	100%

Data source: CRC checkout logs. Items marked n/a were not available in Year 1.

Literacy Coaching Support

In Year 2, RBS and the team of six literacy coaches jointly developed a coaching daily activity sheet (CDAL) to be used to record coaching tasks. The sheet included twelve categories of tasks that the coaches might typically perform, such as conducting observations or meeting with teachers. During the 2007–2008 school year, coaches recorded tasks completed using the sheet, and RBS coded the information using SPSS (see Appendix N-9 for the report and CDAL instrument). The number of logs submitted to evaluators was high: logs submitted by five of six of the coaches represented between 71 and 86 percent of a 190-day year (the school year and an intense period of activity prior to the start of school). The sixth coach’s logs represented 52 percent of the work period. The Year 2 system of logging tasks was an improvement from Year 1 where coaches recorded tasks they had completed only once per month.

RBS entered a total of 5,791 individual tasks from 847 daily activity logs submitted by coaches. Table 33 summarizes the types of activities logged during Year 2. Administrative tasks (N = 1,569) accounted for 27.1 percent of the 5,791 tasks logged, followed by activities

related to training or meeting with teachers (22%), and participation in coach professional development (11.7%). It is important to note that for every interaction between coach and teacher of “substance,” there are corresponding administrative tasks. Professional development for the literacy coaches in Year 2 included, but was not limited to the following: *READ 180* trainings (e.g., Enterprise Edition, Scholastic RED facilitator, and data collection tool training), Santa Cruz Mentor training, and sessions on differentiated instruction. Data provided separately by the Principal Investigator on the professional development opportunities provided to coaches corroborates the information in the coaches’ daily logs. In all, there were 119 hours of professional development offered to coaches, an additional 64 hours of mentorship provided to new Year 2 coaches, and a two-day middle school conference that one coach attended with her school leadership team.

Coaches’ *READ 180* Tasks

In Year 2, coaches logged 600 tasks related to the *READ 180* targeted intervention, accounting for 10.4 percent of the total 5,791 activities documented. The percentage of time devoted to *READ 180* tasks changed little from Year 1, when *READ 180* activities comprised approximately 12.5 percent of all completed tasks (447 of 1,804). *READ 180* tasks included observing *READ 180* teachers or providing them with materials and attending meetings and training sessions related to the program.

Table 33: Type of Coaching Task, Years 1 and 2

Type of Task	Year 1 (N=1,804)		Year 2 (N = 5,791)	
	Frequency	Percent	Frequency	Percent
Coach administrative task	291	16.1	1,569	27.1
Trained or met with teachers	523	29.0	1,272	22.0
Coach professional development	210	11.6	675	11.7
Helped teacher prepare for class	194	10.8	511	8.8
Observed teacher	305	16.9	472	8.2
Non-MCLA school tasks	100	5.5	290	5.0
SR Evaluation tasks	68	3.8	277	4.8
Evening course and U of M related	n/a	n/a	236	4.1
MCLA-related school tasks	n/a	n/a	219	3.8
Assisted teacher in other ways during class	41	2.3	183	3.2
Modeled lesson	51	2.8	68	1.2
Videotaped	17	0.9	19	0.3
Other	4	0.2	n/a	n/a
Total	1,804	100	5,791	100.0

Data source: Coaching daily logs, school years 2006–2007 and 2007–2008

Coach Availability

RBS assessed coaching dosage and availability using two methods: (1) surveys administered at the end of the fall and spring semesters, and (2) a review of entries in the coaching logs that contained references to teacher names. The survey asked respondents how many times they

worked with their literacy coaches, used the CRC materials, and participated in professional development other than MCLA. The purpose of the survey was to collect feedback about respondents' experiences in MCLA and the strategies they implemented as a result of their participation. The number of MCLA teachers completing the survey was 62 in the fall and 54 in the spring, which represented 89 and 82 percent, respectively, of teachers who completed MCLA in Year 2.

As early as October 2007, almost 60 percent of respondents reported that they had met with their coaches more than four times. By spring 2008, three-quarters (75.9%) reported that they had met with their coaches more than four times. These figures are corroborated by data in the coaching logs.

RBS also calculated the number of times each MCLA participant appeared by name in the Year 2 coaches' logs and the number of times they were referenced in a substantive or meaningful way (e.g., working with the coach on lesson plans rather than on administrative tasks). Table 34 summarizes the number of MCLA participants by school who appeared in the coaching logs 10 or more times, considered by RBS to represent adequate or "high" coaching dosage as part of the whole-school intervention. Results show that all MCLA completers (100%) at two schools received high levels of coaching assistance, while three-quarters (76.9%) of participants at one school and about one-third (35.7) percent of participants at another school received high levels of assistance. It should be noted that while coaches encourage teachers to collaborate, the level of participation/involvement is the teachers' prerogative; some teachers may not have accepted coaching assistance and therefore received a low level dosage of coaching support. Moreover, the coach's record-keeping style or level of specificity in of documenting teacher-level interactions affected the analysis because tasks described only generally and not by teacher name in the logs were not included in the frequencies reported. As a result, the number of participants reported to have received high levels of coaching is likely to be understated.

Table 34: Number and Percentages of MCLA Participants with High Coaching Dosage, Year 2

	Number of MCLA Participants, Spring 2008	MCLA Participants with High Coaching Dosage*	
		#	%
A. Maceo	13	13	100%
Hamilton	13	13	100%
Riverview	14	5	35.7%
Sherwood	26	20	76.9%
Total	66	51	77.3%

Data source: Coaching daily logs, school Year 2007–2008

Classroom Level MCLA Implementation

Results from the previously described classroom observations conducted in October 2007 and May 2008 provide insight into teachers' implementation of MCLA strategies. The purpose of the 48 fall observations was to document the extent to which control (N=22) and treatment (N=26) school teachers implemented literacy strategies into their content classes, whereas five

spring observations conducted with paired researchers focused on MCLA participants and on ensuring the reliability of data collected using the observation protocol for future waves of data collection. A more detailed report is included in Appendix N-3.

Table 35 summarizes the presence (or absence) of literacy strategies observed during the fall 2007 data collection. Results show that a greater percentage (38.5%) of teachers from MCLA treatment schools used three or more literacy strategies during the observed class, compared with 18.2 percent of teachers from control schools. The more detailed report in Appendix N-3 shows the breakdown of strategy use by MCLA status within MCLA treatment schools, since some teachers observed using strategies in those schools did not participate in the intervention.

Table 35: Literacy Strategy Use by Teachers in Control and Treatment Classrooms in October 2007

	Treatment Classes (N=26)	Control Classes (N=22)
Used no literacy strategies	6 (23.1%)	10 (45.4%)
Used one strategy	8 (30.7%)	3 (13.6%)
Used two strategies	2 (7.6%)	5 (22.7%)
Used three or more strategies	10 (38.5%)	4 (18.2%)

Table 36 presents the specific literacy strategies used by the 20 MCLA and non-MCLA teachers in treatment schools. Each row in the table represents an individual teacher and his or her grade level, content area, and strategies used during the observed lesson. The most common practices among MCLA teachers included teacher read alouds (N=9) and previewing text (N=7), regardless of the content area taught. Two non-MCLA teachers also read aloud during the observation, and three were observed connecting text to students' everyday lives. Although not considered to be an effective literacy strategy, popcorn reading (which involves individual students taking turns reading text aloud) was observed in one MCLA classroom.

Table 36: Types of Literacy Strategies Used by non-MCLA and MCLA Teachers in Treatment Schools, October 2007 (N=20)

	Grade	Content Area	Types of Literacy Strategies Used					
MCLA Teachers	6th	ELA	Previewing text	Choral reading	Preteaching vocabulary	Activating prior knowledge	Context clue	Connecting text
	7th	ELA	Bubble map	Connecting text				
	6th	ELA	Read aloud	Activating prior knowledge	Monitoring understanding			
	8th	ELA	Read aloud	Previewing text	Monitoring understanding	(Popcorn reading)		
	7th	ELA	Word sorts					
	8th	ELA	Read aloud					
	6th	Science	Activating prior knowledge	Student-generated questions	Questioning for purpose			
	7th	Science	Student-generated questions	Previewing text	Read aloud	Monitoring understanding	Connecting text	
	6th	Science	Glossary use					
	7th	Science	Previewing text	Monitoring understanding				
	7th	Science	Monitoring understanding	Previewing text	Question-answer relationship	Activating prior knowledge		
	8th	Science	Read aloud					
	7th	Social S.	Glossary use	Read aloud	Previewing text			
	8th	Social S.	Read aloud	Choral reading	Monitoring understanding	Previewing text	Word sorts	
Non-MCLA Teachers	6th	Math	Choral reading					
	8th	Math	Read aloud					
	6th	Math	Read aloud					
	8th	Social Studies	Connecting text	Choral reading	Context clue	Monitoring understanding		
	8th	Social Studies	Preteaching vocabulary	Connecting text	Etymology			
	7th	Social Studies	Connecting text					

Data source: Fall 2008 classroom observations collected using the MSRP-COP (Feldman and Feighan, 2007).

Table 37 presents the literacy strategies used by the 12 control teachers and shows that some of the strategies they used were promoted by MCLA. For example, four teachers read aloud during class and three used choral reading strategies during the observation.

Table 37: Types of Literacy Strategies Used by Control Teachers, October 2007 (N=12)

Grade	Content Area	Type of Literacy Strategies Used					
8th	ELA	Read aloud	Preteaching vocabulary	Reflection	Activating prior knowledge	Connecting text	
7th	ELA	Student generating questions	Connecting text				
6th	ELA	Bubble map	Choral reading	Connecting text			
8th	Science	Previewing text	Glossary use				
8th	Science	Monitoring understanding					
7th	Science	Glossary use	Frayer model				
6th	Math	Read aloud	Choral reading	Repeated oral reading			
8th	Math	Glossary use	Preteaching vocabulary				
6th	Math	Choral reading					
6th	Social S.	Read aloud	Preteaching vocabulary				
8th	Social S.	Activating prior knowledge	Preteaching vocabulary	Read aloud	Monitoring understanding	Choral reading	(pop-corn reading)
7th	Social S.	Monitoring understanding					

Data source: Fall 2008 classroom observations collected using the MSRP-COP (Feldman and Feighan, 2007).

In May 2008, pairs of researchers observed ten classes taught by MCLA participants completing the final semester of MCLA. Evaluators observed one eighth grade class, six seventh grade classes, and three sixth-grade classes over a three-day period. In half of the observed classes (N=5), students worked on end-of-year tests (students in four classes completed formal assessments, and students in another class reviewed for an upcoming formal assessment). The mean length of the observations was 54 minutes, and the ten classes had a mean of 18 students, ranging from 15 to 26 students. Evaluators observed three classes each of ELA, science, and social studies, and one mathematics class.

A total of 20 classroom observation protocols were completed for ten classes observed. For the purpose of this summary, RBS randomly selected one of the two protocols that were completed for each class. Overall, the findings from the analysis of the spring 2008 observations revealed:

- Five of the 10 classes observed were sparsely equipped while the other five were rich in resources. (Interviews with teachers revealed that several had put away or removed books and materials and reconfigured their classrooms in preparation for the end of school, which ended one week following the observations). Using a four-point scale where a “1” indicates classroom overcrowding, and a “4” indicates adequate space, observers rated rooms as generally spacious (i.e., rating seven of the 10 classes at a level 4, or spacious).

Six in ten classes had desks arranged appropriately for the task. Half (N=5) of the classes had bare walls. In six of the 10 classes, an evaluator recorded very low availability of books.

- The cognitive demand level of observed lessons was low. (The mean score was a rating of “2” on a six-point scale across four time intervals where “1” indicates low demand and a “6” indicates a high level of demand.
- Eight in ten classes had high student engagement levels for at least three of the four time intervals measured.
- Literacy strategies were implemented in half (N=5) of classes observed. Specific strategies used by teachers are presented below in Table 38.

Table 38: Literacy Strategies Used by Observed MCLA Participants, May 2008 (N=5)

Strategies Used	Grades and Content Areas
Connecting text	7 th grade ELA
Monitoring understanding	6 th grade ELA, 7 th grade ELA
Previewing text	7 th grade science, 7 th grade ELA
Bubble map	7 th grade ELA
Activating prior knowledge	6 th grade ELA, 7 th grade ELA
Read aloud	6 th grade ELA, 7 th grade ELA, 7 th grade science
Choral reading	7 th grade ELA
Questioning for purpose	7 th grade ELA
Preteaching vocabulary	7 th grade ELA
Etymology	7 th grade science
Glossary use	7 th grade social studies, 8 th grade science
Context clue	8 th grade science

Summary of Level of Implementation Attained for Whole-School Intervention

In Years 1 and 2 of the MSRP, MCLA developers held 44 evening sessions grouped by content area for a total of 141.5 hours and required teachers to implement 14 lessons that integrated specific literacy practices (i.e., the use of semantic feature maps, oral retelling strategies, and think-pair-share activities). Literacy coaches helped teachers complete the assignments through lesson modeling, debriefing conferences, observations, and a wide range of other general support activities. In addition to coaching assistance, teachers were encouraged to use materials and resources from an on-site curriculum library maintained by the literacy coaches.

RBS tracked attendance at the MCLA evening classes to determine individual and school-wide program participation in the four schools receiving the first two years of the intervention. Teachers who attended the full year of Year 2 classes had very high participation: approximately two-thirds (65.6%) of completers (N=61) attended 80 percent or more of fall and spring classes.

Although course attendance was high among registered teachers, enrollment in MCLA across the four schools varied widely in Years 1 and 2. In fact, 48 percent of eligible content area teachers participated in one school, compared with 74 percent of eligible teachers in another school in fall 2007. RBS assigned an implementation rating to each school using a formula that

takes into account teachers' course attendance and includes the number of eligible teachers who opted not to participate in the program. All eligible content area teachers in the school were assigned one of four numerical ratings depending on how many MCLA professional development sessions they attended in fall 2007 and spring 2008. The ratings are as follows: teachers who attended 25 percent or fewer of the sessions were given a "1," those attending between 26 and 50 percent of the professional development offered were assigned a "2," teachers participating in between 51 and 75 percent of the professional development offered received a "3," and those who attended 76 to 100 percent of the professional development offered were given a rating of "4."

Other implementation ratings were assigned to the four MCLA schools, including a coaching dosage score, principal involvement rating, and use of materials score. The percentage of MCLA teaches with high coach dosage was determined by the number/percentage of teachers who worked with the coaches 10 or more times during the school year according to data provided in their weekly logs. A principal involvement score of "4" was assigned to each school since all principals attended all of fellowship classes and two key MCLA events. The use of materials was rated through calculation of the number of items a teacher checked out of the CRC: a rating of "2" meant that more than 50 percent of MCLA teachers in the school checked out materials, whereas a rating of "1" indicated that fewer than half checked out the materials at least once (one literacy coach did not provide data on the school's CRC use). Lastly, the school's implementation rating is a composite score based on the previous measures. In Year 1, the coach, principal, and materials scores were not calculated and the implementation score was instead based upon attendance in the MCLA teacher's course (thus these cells contain an "n/a" score where no Year 1 data were provided); however, the Year 2 formula takes into account principal fellowship attendance, individual-level coaching contacts, and teacher-level data on the use of CRC materials.

Once the above ratings were tallied, RBS calculated an average score for each MCLA school and assigned it one of four corresponding schoolwide implementation ratings:

- 1 = minimal program implementation
- 1.1 to 2 = low implementation
- 2.1 to 3 = medium implementation
- 3.1 to 4 = high implementation

Table 39 summarizes these implementation ratings as well as the number and percentage of participants in the intervention by school for Years 1 and 2 of the whole-school intervention.

Table 39: Schoolwide MCLA Participation and Implementation Rankings, Years 1 and 2

3	Number Completing MCLA, Fall Semesters	Number of Eligible Teachers in August	Percent of Eligible Teachers Participating in MCLA	Coaching Dosage Score	Course Participation Score	Principal Involvement Rating	Materials Use Rating	School's Implementation Rating (Includes Nonparticipants)
<i>A. Maceo</i>								
Year 1	14	40	34.6%	n/a	n/a	n/a	n/a	1.9 (minimal)
Year 2	16	33	48.5%	4	1.95	4	2	2.9 (medium)
<i>Hamilton</i>								
Year 1	12	29	41.4%	n/a	n/a	n/a	n/a	2.0 (low)
Year 2	15	24	62.5%	4	2.28	4	missing	2.5 (medium)
<i>Riverview</i>								
Year 1	16	19	84.2%	n/a	n/a	n/a	n/a	3.3 (high)
Year 2	14	19	73.7%	2	2.95	4	1	2.5 (medium)
<i>Sherwood</i>								
Year 1	27	44	61.5%	n/a	n/a	n/a	n/a	2.7 (medium)
Year 2	29	39	74.4%	4	2.89	4	2	3.2 (high)

Results indicate that the level of MCLA implementation was higher in Year 2 than in Year 1. For example, in Year 2, 74.4 percent of all eligible teachers at Sherwood attended MCLA classes at least three-quarters of the time. More than three-quarters (76.9%) of teachers there also worked with the coaches ten or more times on substantive tasks. Although the overall level of MCLA implementation increased between Years 1 and 2, the level was high only at Sherwood. It is possible that the presence of two literacy coaches at that school enabled staff to provide more comprehensive services to teachers; however, A. Maceo Walker also had two coaches and received only a medium implementation rating. That implementation level scores decreased at Riverview may be more indicative of the recordkeeping of its coach than any actual level of program implementation. Overall, the level of MCLA implementation ranged from medium to high at the four schools in Year 2.

VI. Evaluation of the Impacts of the Whole-school Intervention: Years 1 and 2

Sample Selection

Prior to the beginning of the 2006–2007 school year, the MSRP schools were grouped into pairs matched on the most recent (spring 2005) TCAP scores in reading/language arts and mathematics. These schools were quite homogenous on other available characteristics, such as

gender, race/ethnicity, students receiving free or reduced-price lunches, ELL students, and students with disabilities. One of the schools in each pair was randomly assigned to participate in MCLA during Years 1 and 2. The other four schools will participate in MCLA in Years 3 and 4.

There were 289 teachers of core content areas (language arts, mathematics, social studies, and science) identified in the eight schools during Year 1. The 145 content area teachers in the four MCLA schools were invited to participate in this two-year professional development program. Part V of this report describes the participation rates of these teachers in Years 1 and 2.

The sample of students for studying the immediate, one-year impacts of the MCLA whole-school intervention consists of those students who were enrolled in participating schools for a majority of the instructional days between September 18, 2006, and May 3, 2007, in Year 1 or sixth-grade students meeting the same criterion between September 7, 2007 and May 13, 2008 in Year 2.¹⁸ The sample for studying the long-term, two-year impacts of MCLA consists of the seventh- and eighth-grade students in Year 2 meeting this criterion for both years. In addition, they had to be enrolled in a school from the same design group, treatment or control, in both years. As in the ITT sample for the analysis of the *READ 180* impacts, this MCLA impact sample excluded students receiving special education services. Table 40 describes the numbers of students in these different samples.

Table 40: Number of Students Enrolled in Striving Reader Study Schools for a Majority of Instructional Days By Year, Grade, and MCLA Design Group

Year	Grade	Students in MCLA Treatment Schools	Students in MCLA Control Schools	Total Number of Students
1	6	690	817	1507
	7	883	945	1828
	8	857	821	1678
	All	2430	2583	5013
2	6	660	779	1439
1 and 2	6→7	471	611	1082
	7→8	577	676	1253
	Both	1048	1287	2335

Data Collection

The measures of teacher outcomes for determining the impact of MCLA on core content area teachers are two composite indices based on 24 items from a teacher survey administered at the beginning and end of the 2006–2007 school year and at the end of the 2007–2008 school year (see Appendix N-2B). The first index measures how prepared the teachers report they are to engage in each of the 24 literacy activities; the second measures the frequency that they report en-

¹⁸ The September dates represent the point when school enrollments were considered stable enough to identify and randomly assign the eligible pool of struggling readers to the targeted intervention, *READ 180*. The May dates represent the end-of-year administration of the ITBS.

gaging in each of these activities. Preparedness and frequency items were measured with 5-point scales. Both composite indices are the average rating assigned to the 24 items.¹⁹

The measures of student outcomes for determining the impact of MCLA on students are the ITBS and the TCAP. (Details about these measures are provided on page 15.) The ITBS was administered twice during Year 1—during the week beginning September 18, 2006, and during the week beginning April 30, 2007—by classroom teachers to all students in the MSRP schools, except those in self-contained special education classrooms and a very small number whose parent did not consent to the student’s participation in the testing. The spring 2007 test scores measured treatment and control student reading achievement levels at the end of Year 1. The fall 2006 test scores were used to control for random differences in reading achievement levels between treatment and control students at the beginning of the year, as well as reduce the within-school error variance in the spring 2007 test scores.

The ITBS was administered twice also during Year 2—during the weeks beginning September 17, 2007, and May 12, 2008—by classroom teachers in MSRP schools. The fall administration was only for students in sixth grade; all students in MSRP schools took the spring administration. The spring 2008 test scores measured treatment and control student reading achievement levels at the end of Year 2. The fall 2007 test scores were used as control variables for cross-sectional, immediate impact analyses of Year 2 achievement for sixth-grade students. The fall 2006 test scores were used as control variables for cross-sectional, long-term impact analyses of Year 2 achievement for students in the seventh and eighth grades, and as baseline test scores for longitudinal analyses of growth in achievement over the two years for students in the seventh and eighth grades.

The TCAP is administered by MCS for the state on or about the first week in April each year. The spring 2007 test scores measured treatment and control student achievement levels in the four core content areas at the end of Year 1. The spring 2006 scores in the same content area were used to control for random treatment-control differences and reduce within-school error variance in cross-sectional, immediate impact analyses of spring 2007 scores in Year 1. The spring 2007 test scores were used as control variables for cross-sectional, immediate impact analyses of Year 2 achievement for sixth-grade students. The spring 2006 test scores were used as control variables for cross-sectional, long-term impact analyses of Year 2 achievement for students in the seventh and eighth grades, and as baseline test scores for longitudinal analyses of growth in achievement over the two years for students in the seventh and eighth grades.

Data Analysis

Cross-sectional impact analyses of teacher outcomes, their preparedness to use and frequency of use of 24 literacy activities, were conducted to assess the immediate effects of first-year participation in MCLA and the long-term effects of two years of participation on these two measures. Reliability tests of these items using Year 1 baseline survey data produced a Cronbach's alpha of .961 for the preparedness items and .947 for the items that measure frequency of strat-

¹⁹ Factor analysis was used to study the structure underlying responses to the preparedness and frequency items. The results support the use of a single composite of all 24 items, for both preparedness and frequency.

egy use (full details about these analyses are included as Appendix P). These analyses compared the average preparedness and frequency indices obtained from the core content area teachers in MCLA treatment and control schools at the end of Years 1 and 2. Multi-level regression analysis models were used to estimate and test the statistical significance of the difference between the preparedness and frequency indices of teachers in MCLA treatment and control schools. Two-level models were employed that express the year-end indices as a function of teacher and school variables. The MCLA treatment variable was included at the school level of these models. The complete specification of the multi-level regression model employed to determine the immediate and long-term impacts of the MCLA intervention on teacher outcomes is provided in Appendix G.

The preparedness and frequency indices based on the year-end administrations of the teacher survey were the dependent variables in these analyses. The two indices based on the baseline, beginning-of-Year 1 administration of the teacher survey—representing the same index as the dependent variable—were included as the main teacher-level covariates. Other control variables at the teacher and school level were tested for inclusion as covariates in these analyses. Table 41 summarizes the dependent and independent variables and the covariates included in the analyses of teacher outcomes.

Cross-sectional impact analyses of student achievement in reading and the four core content areas were conducted to assess the immediate effects of first-year participation of schools in MCLA on student outcomes for the 5,013 students in the sixth, seventh, and eighth grades who were enrolled in an MSRP school for a majority of instructional days in Year 1. Also, to investigate any suggestions of interactions of MCLA impact and grade level, separate analyses were conducted for students in the sixth, seventh, and eighth grades. A separate cross-sectional analysis of the 1,439 students in sixth grade who were enrolled in an MSRP school for a majority of instructional days in Year 2 was conducted to determine whether the immediate impact of MCLA varied between Years 1 and 2.

Similar cross-sectional analyses were conducted to assess the long-term, two-year effects of MCLA on student achievement at the end of Year 2 for the 2,335 students in the seventh and eighth grades who were enrolled in an MSRP school in the same MCLA design group in Years 1 and 2. In addition, the long-term effects of MCLA on growth in student achievement during Years 1 and 2 were examined using longitudinal analyses of the achievement of these 2,335 students at the beginning and end of Year 1 and the end of Year 2.

Multi-level regression analysis models were used to estimate and test the statistical significance of the difference between the reading and content area achievement of students in MCLA treatment and control schools. Two-level models were employed for the Year 1 cross-sectional analyses that express the spring ITBS and TCAP scores as a function of student and school variables.²⁰ The spring 2007 ITBS and TCAP scores were the dependent variables. The 2006 ITBS

²⁰ Three-level models, employing school, teacher, and student variables were explored. These analyses proved to be relatively complex and equivocal due to each student's having different teachers for the core content areas and significant amounts of missing teacher data. Also, the results did not vary noticeably from the results of the two-level models. The evaluation team decided to omit these models from the impact analyses.

and TCAP scores—representing the same test or subject as the dependent variable—were included as the main student-level covariate. Other control variables at the student and school level were tested for inclusion as covariates in these analyses. The MCLA treatment variable was included at the school level of these models.

Table 41: All Variables Included in MCLA Impact on Teacher Outcomes—Analytical Models for Years 1 and 2

Variable	Level	Coding / Range
Dependent		
Year-End Preparedness Index (2007/2008)*	Teacher	1-5; Not at All; A Little; Prepared; Well Prepared; Could Teach Others
Year-End Frequency Index (2007/2008)*	Teacher	1-5; Never; Rarely; Sometimes; Often; Almost Always
Independent		
MCLA Participation	School	Yes = 1; No = 0
Covariates (Fall 2006)		
Baseline Preparedness Index	Teacher	1-5; 5 represents highest preparedness
Baseline Frequency Index	Teacher	1-5; 5 represents highest frequency
English Language Arts Teacher	Teacher	Yes = 1; No = 0
Age Level	Teacher	1-6: 20's; 30's; 40's; 50's; 60's; 70's
Gender	Teacher	Female = 1; Male = 0
African-American	Teacher	Yes = 1; No = 0
Masters Degree or Higher	Teacher	Yes = 1; No = 0
Licensed in Grade/Subject Taught	Teacher	Yes = 1; No = 0
Prof Dev in Integrating Literacy in Class	Teacher	1-4: None; 1-8 hrs; 9-32 hrs; 32+ hrs
Years Full-time Teacher	Teacher	1-7: Never; 0-2; 3-5; 6-10; 11-20; 21-30; 30+
Years Full-time at Current School	Teacher	1-7: Never; 0-2; 3-5; 6-10; 11-20; 21-30; 30+
Percentage ELA Teachers	School	0-100
Avg Age Level	School	1-6: 20's; 30's; 40's; 50's; 60's; 70's
Percentage Female Teachers	School	0-100
Percentage African-American Teachers	School	0-100
Percentage Masters Degree or Higher	School	0-100
Percentage Licensed in Grade/Subject	School	0-100
Avg PD in Integrating Literacy in Class	School	1-4: None; 1-8 hrs; 9-32 hrs; 32+ hrs
Avg Years Teachers Full-time	School	1-7: Never; 0-2; 3-5; 6-10; 11-20; 21-30; 30+
Avg Years Teachers Full-time at School	School	1-7: Never; 0-2; 3-5; 6-10; 11-20; 21-30; 30+

* Second date applies for analyses in Year 2.

Similar two-level models were employed for the cross-sectional analyses of student achievement at the end of Year 2. The only differences were the use of spring 2008 ITBS and TCAP test scores as the dependent variables and, for the sixth-grade students, the use of 2007 ITBS and TCAP test scores as one of the student covariates.

Table 42: All Variables Included in MCLA Impact—Analytical Models for Years 1 and 2

Variable	Level	Coding / Range
Dependent		
Spring 2007/2008 ITBS Total Reading*	Student	Standard Score 100-350
Spring 2007/2008 ITBS Comprehension*	Student	Standard Score 100-350
Spring 2007/2008 ITBS Vocabulary*	Student	Standard Score 100-350
Spring 2007/2008 TCAP Reading/LA*	Student	Scale Score 300-750
Spring 2007/2008 TCAP Mathematics*	Student	Scale Score 300-750
Spring 2007/2008 TCAP Science*	Student	Scale Score 100-300
Spring 2007/2008 TCAP Social Studies*	Student	Scale Score 100-300
Independent		
MCLA Participation	School	Yes = 1; No = 0
Covariates		
Test Score at End of Year 1	Time	Yes = 1; No = 0
Test Score at End of Year 2	Time	Yes = 1; No = 0
Fall 2006/2007 ITBS Total Reading** ***	Student	Standard Score 100-350
Fall 2006/2007 ITBS Comprehension** ***	Student	Standard Score 100-350
Fall 2006/2007 ITBS Vocabulary** ***	Student	Standard Score 100-350
Spring 2006/2007 TCAP Reading/LA** ***	Student	Scale Score 300-750
Spring 2006/2007 TCAP Mathematics** ***	Student	Scale Score 300-750
Spring 2006/2007 TCAP Science** ***	Student	Scale Score 100-300
Spring 2006/2007 TCAP Social Studies** ***	Student	Scale Score 100-300
Gender	Student	Female = 1; Male = 0
African-American	Student	Yes = 1; No = 0
Hispanic	Student	Yes = 1; No = 0
Free or Reduced Lunch (2006/2007)**	Student	Yes = 1; No = 0
English Language Learner (2006/2007)**	Student	Yes = 1; No = 0
Enrolled in Grade 7 in Year 1	Student	Yes = 1; No = 0
Enrolled in Grade 8 in Year 1/Year 2*	Student	Yes = 1; No = 0
Percentage Female (Fall 2006)	School	0-100
Percentage African-American (Fall 2006)	School	0-100
Percentage Special Ed (Fall 2006)	School	0-100
Percentage FRL (Fall 2006)	School	0-100
Percentage ELL (Fall 2006)	School	0-100
School Enrollment (Fall 2006)	School	400-1200

* *Second date applies for analyses in Year 2.*

** *Only used in cross-sectional, not in longitudinal analyses.*

*** *Second date applies for Grade 6 analyses in Year 2.*

Three-level, longitudinal models were used to estimate and test the statistical significance of the difference between growth in reading and content area achievement over the two years for students in MCLA treatment and control schools. The first level expresses each student's test scores at three points in time as the sum of the baseline score plus the growth at the end of Year 1 and the additional growth at the end of Year 2. The second level expresses the student's baseline test score as a function of student characteristics. The third level expresses the average base-

line score as a function of MCLA participation and other school covariates, and expresses the Year 1 and Year 2 growth in test scores as a function of MCLA participation.

The complete specification of the multi-level regression models employed to determine the immediate and long-term impacts of the MCLA whole-school intervention is provided in Appendix I. Table 42 summarizes the dependent and independent variables and the covariates included in these analyses.

Selection of Covariates

There are different approaches to including/excluding covariates in multi-level regression, as there are in single-level regression analyses. The approach that was used in the analyses of teacher outcomes was to (1) include all teacher covariates and MCLA school-level variables in the model, (2) run the model, (3) add the school-level covariate with highest potential for accounting for between school variance, (4) keep the covariate if its significance level was less than 0.2, (5) repeat steps 2 to 4 until no more covariates with p-values less than 0.2 could be added, (6) run the model, (7) eliminate the teacher covariate with the lowest significance level (highest p-value) not less than 0.2, and (8) repeat steps 6 and 7 for the rest of the teacher covariates until the remaining covariates had p-values less than 0.2.²¹ The approach used in the analysis of student outcomes was to (1) include all student- and school-level covariates in the model, (2) run the model, (3) eliminate the school covariate with the lowest significance level (highest p-value) not less than 0.2, (4) repeat steps 2 and 3 until the remaining covariates had p-values less than 0.2, and (5) repeat steps 2 to 4 for the student covariates.

Treatment of Missing Data

Procedures for imputing missing values were not employed.

Description of Samples of Teachers and Students for MCLA Impact Analyses

Equivalence on Teacher Characteristics

As described above, 289 teachers of core content areas were identified in the eight participating schools. Four of the schools were randomly assigned to the MCLA treatment and the other four to the control group. The percentages of teachers teaching each content area and possessing other characteristics are presented in Table 43.

The percentages for teachers in control and treatment schools are quite similar. The control school teachers appear to have a little more experience. Including these characteristics as teacher-level covariates in the analytical models helped to control for these small differences, as well as reduce the within-school error variance in the dependent variables.

²¹ This procedure varies from the one employed for models of student outcomes due to the large number of school-level covariates.

Table 43: Selected Characteristics of the Teacher Sample for MCLA Impact Analyses

Teacher Characteristic	Control ^a	MCLA ^a	Total ^a
Teaches Language Arts	32.1%	37.5%	34.8%
Teaches Mathematics	20.1%	19.1%	19.6%
Teaches Science	17.9%	18.4%	18.1%
Teaches Social Studies	19.4%	20.6%	20.0%
Female	74.2%	74.2%	74.2%
Male	25.8%	25.8%	25.8%
African-American	86.7%	88.0%	87.4%
Masters Degree or Higher	53.9%	59.8%	56.9%
Licensed in Grade/Subject Taught	85.4%	79.3%	82.3%
PD in Integrating Literacy in Class (more than 8 hrs in past 12 months)	44.2%	39.5%	41.9%
Years Full-time Teacher (more than 5 years)	67.8%	57.6%	62.6%
Years Full-time at Current School (more than 5 years)	14.4%	13.3%	13.9%
Years Full-time in Memphis (more than 5 years)	52.2%	44.4%	48.4%

^a These percentages are based on different numbers of teachers due to variations in response rates to different items on the teacher survey.

Equivalence on Baseline Teacher Indices

Comparisons between teachers in treatment and control schools on the baseline indices of preparedness to use and frequency of use of the 24 literacy activities were carried out. Table 44 describes these differences. With random assignment of schools, the teachers in treatment and control schools should be similar on both indices.

Table 44: Comparison of Teachers in MCLA Treatment and Control Schools on Baseline Indices

Index Score	Means		Signif. Level
	Control	MCLA	
Preparedness Index	3.52 (89) ^a	3.54 (81)	0.877
Frequency Index	3.73 (88)	3.62 (72)	0.315

^a Numbers in parentheses are the numbers of teachers in each group having a valid index score.

For both indices, the estimated difference between teachers in control and treatment schools is not statistically significant ($p < 0.05$). It is appropriate to conclude that the MCLA control and treatment school teachers were statistically equivalent on both indices at the beginning of the school year.

Equivalence on Student Characteristics

As described above, 5,013 students were enrolled in MSRP schools for a majority of instructional days between September 18, 2006, and May 3, 2007, in Year 1 of the MSRP study. The grade level and other demographic characteristics of these students are presented in Table 45.

Table 45: Demographic Characteristics of the Year 1 MCLA Student Sample

Student Characteristic	Control ^a	MCLA ^a	All Schools ^a
Enrolled in Grade 6	817 (31.6%)	690 (28.4%)	1507 (30.1%)
Enrolled in Grade 7	945 (36.6%)	883 (36.3%)	1828 (36.5%)
Enrolled in Grade 8	821 (31.8%)	857 (35.3%)	1678 (33.5%)
Female	1295 (50.1%)	1291 (53.1%)	2586 (51.6%)
Male	1288 (49.9%)	1139 (46.9%)	2427 (48.4%)
African-American	2375 (91.9%)	2374 (97.7%)	4749 (94.7%)
Hispanic	193 (7.5%)	49 (2.0%)	242 (4.8%)
Free or Reduced Lunch	2235 (86.5%)	2175 (89.5%)	4410 (88.0%)
English Language Learner	143 (5.5%)	27 (1.1%)	170 (3.4%)
Total	2583	2430	5013

^a Percentages are based on the total number of students in control, treatment, or all schools.

All but 22 students were either African-American or Hispanic, supporting the creation of two dichotomous covariates to represent membership in these two race/ethnicity groups. The differences in demographic composition of the treatment and control groups were relatively minor, although all were statistically significant ($p < .05$) given the large number of students overall. Including these characteristics as student-level covariates in the analytical models helped to control for these small differences, as well as reduce the within-school error variance in the dependent variables.

In Year 2, 1,439 sixth-grade students were enrolled in MSRP schools for a majority of instructional days between September 7, 2007, and May 13, 2008. The demographic characteristics of these students are presented in Table 46.

Table 46: Demographic Characteristics of the Year 2 Grade 6 MCLA Student Sample

Student Characteristic	Control ^a	MCLA ^a	All Schools ^a
Female	390 (50.1%)	314 (47.6%)	704 (48.9%)
Male	389 (49.9%)	346 (52.4%)	735 (51.1%)
African-American	702 (90.1%)	650 (98.5%)	1352 (94.0%)
Hispanic	67 (8.6%)	6 (0.9%)	73 (5.1%)
Free or Reduced Lunch	711 (91.3%)	629 (95.3%)	1340 (93.1%)
English Language Learner	55 (7.1%)	6 (0.9%)	61 (4.2%)
Total	779	660	1439

^a Percentages are based on the total number of students in control, treatment, or all schools.

The differences between students in MCLA treatment and control schools were relatively small, but statistically significant ($p < .05$) for all characteristics except gender. Treatment schools had higher percentages of African-American students and students receiving free or reduced-price lunch. They had lower percentages of Hispanic and ELL students. These characteristics are included in analyses to help control for these differences.

Also in Year 2, 2,335 students in the seventh and eighth grades remained enrolled in an MSRP school in the same MCLA design group for a majority of instructional days out of the

3,335 students meeting that criterion in grades 6 and 7 in Year 1. The demographic characteristics of these “stayers,” compared with the 1,171 “leavers”²² who were not in a participating school on May 13, 2008, are presented in Table 47.

Table 47: Demographic Characteristics of the Year 2 “Stayers” and “Leavers” from the Year 1 MCLA Sample

Student Characteristic	“Stayers”		“Leavers”	
	Control ^a	MCLA ^a	Control ^a	MCLA ^a
Enrolled in Grade 6 → 7	611 (47.5%)	471 (44.9%)	210 (43.5%)	234 (42.8%)
Enrolled in Grade 7 → 8	676 (52.5%)	577 (55.1%)	270 (55.9%)	313 (57.2%)
Female	655 (50.9%)	555 (53.0%)	222 (46.0%)	277 (50.6%)
Male	632 (49.1%)	493 (47.0%)	261 (54.0%)	270 (49.4%)
African-American	1182 (91.8%)	1032 (98.5%)	442 (91.5%)	524 (95.8%)
Hispanic	100 (7.8%)	15 (1.4%)	37 (7.7%)	19 (3.5%)
Free or Reduced Lunch	1138 (88.4%)	942 (89.9%)	429 (88.8%)	505 (92.3%)
English Language Learner	79 (6.1%)	11 (1.0%)	23 (4.8%)	10 (1.8%)
Total	1287	1048	483	547

^a Percentages are based on the total for the control and treatment groups for each type of student.

Differences between treatment and control groups for the 2,335 seventh- and eighth-grade students in Year 2 who remained in the MCLA sample are relatively small. There are higher percentages of females, African-Americans, and students receiving free or reduced-price lunch in the treatment group, and higher percentages of Hispanic and ELL students in the control group. Overall, differences in treatment and control groups are very similar for those who remained and those who did not.

Equivalence on Baseline Student Achievement

Comparisons between students enrolled (for a majority of instructional days) in MCLA treatment and control schools on the baseline 2006 ITBS and TCAP test scores were carried out for the 5,013 students in Year 1 and the 2,335 “stayers” in the seventh and eighth grades in Year 2. Comparisons were also carried out on the baseline 2007 test scores for the 1,439 students in sixth grade in Year 2.

Table 48 describes the differences between Year 1 students in MCLA treatment and control schools on baseline 2006 test scores for the three ITBS standard scores and the four TCAP content area assessments. With random assignment of schools, the students in treatment and control schools should be similar on all seven test scores.

²² The 1,171 “leavers” included 127 students who were enrolled in an MSRP school on 5/13/08, but the school was in a different MCLA design group.

Table 48: Comparison of Year 1 Students in MCLA Treatment and Control Schools on Baseline 2006 Scores on Each Achievement Test

Test Score	Means		Signif. Level
	Control	MCLA	
ITBS Total Reading Standard Score	205.7 (2235)	204.3 (2119)	0.045
ITBS Comprehension Standard Score	203.8 (2240)	203.3 (2133)	0.521
ITBS Vocabulary Standard Score	207.5 (2244)	205.3 (2129)	0.003
TCAP Reading/LA Scale Score	502.2 (2350)	502.9 (2294)	0.492
TCAP Mathematics Scale Score	505.4 (2347)	502.9 (2293)	0.015
TCAP Science Scale Score	187.7 (2308)	190.2 (2285)	0.000
TCAP Social Studies Scale Score	193.0 (2312)	192.0 (2278)	0.048

^a Numbers in parentheses are the numbers of students in each group having a valid test score.

Students in control schools were significantly ($p < 0.05$) higher on four of the measures: ITBS Total Reading and Vocabulary and TCAP Mathematics and Social Studies. Students in treatment schools scored significantly higher on TCAP Science. Given the small number of schools randomly assigned, the use of 2005 TCAP scores in creating matched pairs of schools, the large number of students, and the exclusion of students who did not attend participating schools for a majority of instructional days, these differences are not too disturbing. Treating the 2006 test scores as covariates in the analyses of the impact of MCLA on 2007 test scores allows some adjustment to be made for these differences. Very similar results were obtained when only the eligible struggling readers (i.e., as described on pages 6 – 7, those who were defined as struggling for the targeted intervention) were included in these comparisons.

Table 49 describes the differences between the Year 2 sixth-grade students in MCLA treatment and control schools on the baseline 2007 test scores.

Table 49: Comparison of Year 2 Grade 6 Students in MCLA Treatment and Control Schools on Baseline 2007 Scores on Each Achievement Test

Test Score	Means		Signif. Level
	Control	MCLA	
ITBS Total Reading Standard Score	196.6 (658)	192.1 (544)	0.000
ITBS Comprehension Standard Score	196.5 (658)	192.0 (545)	0.001
ITBS Vocabulary Standard Score	196.8 (659)	192.1 (549)	0.000
TCAP Reading/LA Scale Score	504.4 (725)	500.0 (622)	0.004
TCAP Mathematics Scale Score	502.9 (728)	498.6 (622)	0.008
TCAP Science Scale Score	193.2 (710)	190.5 (622)	0.005
TCAP Social Studies Scale Score	197.6 (710)	194.6 (621)	0.000

^a Numbers in parentheses are the numbers of students in each group having a valid test score.

Students in the control schools performed significantly higher on all seven test scores. Treating the 2007 test scores as covariates in the analyses of the impact of MCLA on 2008 test scores allows some adjustment to be made for these differences. A similar analysis for eligible struggling readers in this sample yielded no significant differences on any of the seven measures. Table 50 describes the 2006 baseline test score differences between students in treatment and control schools who were enrolled in an MSRP school in the same design group for a majority of instructional days in both Years 1 and 2, i.e., the 2,335 students in Year 2 in seventh and eighth grade who “stayed.”

Table 50: Comparison of Year 2 “Stayers” in MCLA Treatment and Control Schools on Baseline 2006 Scores on Each Achievement Test

Test Score	Means		Signif. Level
	Control	MCLA	
ITBS Total Reading Standard Score	201.2 (1157)	199.7 (955)	0.013
ITBS Comprehension Standard Score	199.5 (1160)	198.0 (955)	0.170
ITBS Vocabulary Standard Score	202.7 (1163)	199.7 (955)	0.003
TCAP Reading/LA Scale Score	500.2 (1212)	499.0 (1025)	0.385
TCAP Mathematics Scale Score	502.1 (1210)	500.8 (1025)	0.309
TCAP Science Scale Score	189.3 (1182)	191.2 (1022)	0.007
TCAP Social Studies Scale Score	194.2 (1185)	193.9 (1020)	0.732

^a Numbers in parentheses are the numbers of students in each group having a valid test score.

The students in control schools scored significantly higher on the ITBS Total Reading and Vocabulary measures, and significantly lower on the TCAP Science assessment. These differences also appear very similar to the treatment-control differences for the Year 1 sample. Two-way ANOVAs were used to test for an interactive effect of “staying” and treatment/control school on 2006 test scores. No significant interactions were found.

Impact of MCLA on Teacher Preparedness and Frequency of Use in Years 1 and 2

Immediate Impact of MCLA in Year 1

Table 51 summarizes the results of the analysis of the immediate impact of MCLA participation on teacher preparedness to use and frequency of use of 24 literacy activities at the end of the first year of participation. (The complete results of the multi-level analyses of the MCLA impact on these two indices can be found in Appendix H in Tables H1 and H2.) A positive impact means the teachers in MCLA treatment schools averaged higher preparedness or frequency of use than teachers in MCLA control schools, controlling for covariates included in the final analytical model (see Tables H1 and H2). A negative impact means teachers in the control schools averaged higher than teachers in the treatment schools.

The estimated impacts of MCLA on these two teacher outcomes are large relative to the 1–5 scale used by the two indices. Both were statistically significant ($p < 0.05$) and the effect sizes were in the moderate range. Clearly, these results support the conclusion that the MCLA professional development had a significant positive impact on teachers’ reports of their preparedness to use a variety of literacy activities and on the frequency with which they use these activities.

Table 51: Immediate Impact of MCLA Treatment and Control Schools on Year-End Indices of Teacher Preparedness and Frequency of Use

Test Score	Unadjusted Means		Adjusted Means		Est. Impact	Effect Size	Signif. Level
	Control	MCLA	Control	MCLA			
Preparedness Index	3.57 (49) ^a	3.92 (49)	3.52	3.93	0.41	0.75	0.012
Frequency Index	3.69 (49)	3.93 (43)	3.64	4.00	0.36	0.61	0.022

^a Numbers in parentheses are the numbers of teachers in each group having valid index scores from the baseline 2006 administration and the spring 2007 administration.

Long-Term Impact of MCLA in Year 2

Table 52 summarizes the results of the long-term, two-year impact of MCLA participation on teacher preparedness to use and frequency of use of 24 literacy activities at the end of the second year of participation. (The complete results of the multi-level analyses of the MCLA impact on these two indices can be found in Appendix H in Tables H3 and H4.)

Table 52: Long-Term (Two-Year) Impact of MCLA Treatment and Control Schools on End-of-Year 2 Indices of Teacher Preparedness and Frequency of Use

Test Score	Unadjusted Means		Adjusted Means		Est. Impact	Effect Size	Signif. Level
	Control	MCLA	Control	MCLA			
Preparedness Index	3.68 (51) ^a	4.04 (46)	3.68	4.10	0.42	0.61	0.048
Frequency Index	3.82 (52)	3.98 (47)	3.93	4.04	0.11	0.22	0.359

^a Numbers in parentheses are the numbers of teachers in each group having valid index scores from the baseline 2006 administration and the spring 2007 administration.

The estimated MCLA impact on the preparedness index at the end of Year 2 was again substantial and statistically significant ($p < 0.05$) and the effect size was moderately large. However, the long-term impact on the frequency of use index was not statistically significant and the effect size was at the low end of the conventional small range.

Differential Attrition

Similar to the finding reported earlier that many students with baseline 2006 test scores did not have spring 2007 scores, many of the teachers who had preparedness and frequency index scores based on the baseline teacher survey did not have index scores from the year-end administration of the survey. Therefore, it is important to examine whether the teachers from MCLA treatment schools who were missing year-end scores were different from the teachers from MCLA control schools who were missing these scores. If they were different, one could argue that the above estimated impacts were biased. That is, the treatment schools may have lost teachers that would have scored higher (or lower) than the teachers lost from the control schools.

This potential differential attrition was studied by comparing the average baseline 2006 index scores of the teachers who also had a year-end 2007 score to the average of all teachers with

baseline 2006 index scores, the difference being attributable to the attrition of teachers. This comparison was done for the teachers in both the treatment and control schools.²³ If this attrition effect were higher/lower in one design group of schools, this differential attrition would have to be acknowledged as possibly biasing the estimated impact of MCLA.

In summary, the effects of attrition in both treatment and control schools on baseline 2006 index scores were quite small and not statistically significant ($p < 0.05$). Separate analyses for the attrition between the 2006 baseline and the spring 2008 results were not carried out since the teachers with spring 2007 index scores were virtually the same as those with spring 2008 index scores. It seems reasonable to conclude that differential attrition was not a biasing factor affecting the interpretation of the estimated impacts of MCLA.

Conclusions

The significant immediate impacts of the MCLA schoolwide intervention on these teacher outcomes support the conclusion that teachers' perceptions of their preparation to use and frequency of use of a variety of literacy activities were improved more in the MCLA treatment schools at the end of Year 1. At the end of Year 2, the long-term impact of MCLA was maintained for teachers' perceptions of their preparation to use literacy activities. However, it was not maintained for frequency of use. Acknowledging the subjectivity of the teacher outcome measures and the possibility of a "Hawthorne Effect" (teachers gave themselves higher marks simply because they knew they were receiving special treatment), these results provide some support for the validity of the logic model for the whole-school intervention, i.e., teachers will improve their literacy instruction, which will lead to better student achievement.

Further Analyses

Exploratory analyses are planned that will study the relationship between MCLA impact and the amount of teacher participation in the MCLA professional development. A greater impact for teachers with higher levels of participation would provide further support for the effectiveness of the whole-school intervention. On the other hand, if non-participating teachers in MCLA treatment schools demonstrate as much or more preparation for and/or frequency of use of literacy activities, this would suggest that other factors are contributing to these teacher outcomes.

Impact of MCLA Participation on Student Achievement in Years 1 and 2

Immediate Impact of MCLA in Year 1

Table 53 summarizes the results of the analysis of the immediate impact of MCLA participation on student reading achievement measured by the ITBS and student achievement in the four core content areas measured by the TCAP. (The complete results of the multi-level analyses of the MCLA impact on these seven test scores can be found in Appendix J in Tables J1 – J7.)²⁴ The

²³ The statistical analysis was the same as that used to study differential attrition of students reported earlier, viz., a univariate 2x2 factorial ANOVA.

²⁴ Note the consistently low levels of between-school variance in these tables as indicated by very small interclass correlation coefficients. This outcome led to a decision to not investigate whether school characteristics moderated estimated impacts.

estimated impact is the difference between the adjusted means for students in MCLA treatment schools and students in control schools. A positive impact means the students in MCLA treatment schools averaged higher achievement on the particular test than the students in control schools, controlling for covariates included in the final analytical model (see Appendix J1 – J7). A negative impact means the students in control schools averaged higher than those in treatment schools.

Table 53: Immediate Year 1 MCLA Impact on Spring 2007 Scores on Each Achievement Test

Test Score	Unadjusted Means		Adjusted Means		Est. Impact	Effect Size	Signif. Level
	Control	MCLA	Control	MCLA			
ITBS Total Reading Standard Score	208.8 (1925) ^a	208.8 (1831)	207.8	207.6	-0.2	0.01	0.900
ITBS Comprehension Standard Score	205.7 (1932)	205.8 (1835)	202.9	207.1	4.2	0.15	0.067
ITBS Vocabulary Standard Score	211.8 (1938)	210.2 (1854)	211.8	208.9	-2.9	0.13	0.125
TCAP Reading/LA Scale Score	517.0 (2301)	515.1 (2240)	519.3	513.6	-5.7	0.19	0.000
TCAP Mathematics Scale Score	522.4 (2297)	515.1 (2240)	521.2	515.1	-6.1	0.18	0.061
TCAP Science Scale Score	192.2 (2212)	193.1 (2222)	193.1	192.0	-1.1	0.07	0.355
TCAP Social Studies Scale Score	193.5 (2205)	191.4 (2212)	193.2	191.3	-1.9	0.13	0.337

^a Numbers in parentheses are the numbers of students in each group having valid test scores from the baseline 2006 administrations and the spring 2007 administrations.

The estimated MCLA impacts vary across the seven test scores—from over four standard score points on the ITBS comprehension subtest in favor of the students in treatment schools to more than six scale score points on the TCAP Mathematics assessment in favor of the students in control schools. Only one estimated impact was statistically significant ($p < 0.05$)—on the TCAP Reading/LA assessment favoring the control students. But the effect size was less than 0.2, as was the case for all of the other estimated impacts. Overall, the comparisons favored the students in control schools. However, with one exception, there is no reason to reject the hypothesis that the average achievement of all students in the treatment and control schools was at the same levels at the end of Year 1. These impact analyses were also carried out for the eligible struggling readers. None of the estimated impacts were statistically significant for struggling readers (see Tables J8 – J14).

Immediate Impact of MCLA at Each Grade in Year 1

The analyses of the immediate impact of MCLA in Year 1 reported above were based on the total sample of students in the sixth to eighth grades who were enrolled in a Striving Readers school for the majority of instructional days. The same analyses were also carried out separately for students in each of these three grades. The results of analyses for sixth grade can be compared with results for sixth grade in Years 2 to 4 to see if there are any changes in the immediate impact of MCLA over time. The results for seventh- and eighth-grade students in Year 1 allow a comparison of the impact of this whole-school intervention across grades. Table 54 presents the results of the analyses of the immediate impact of MCLA on students in sixth grade in Year 1.

Table 54: Immediate Year 1 Impact of MCLA on Spring 2007 Scores on Each Achievement Test—Grade 6

Test Score	Unadjusted Means		Adjusted Means		Est. Impact	Effect Size	Signif. Level
	Control	Treatment	Control	Treatment			
ITBS Total Reading Standard Score	197.4 (710) ^a	196.6 (585)	194.8	198.3	3.5	0.17	0.069
ITBS Comprehension Standard Score	194.0 (710)	196.2 (585)	191.6	197.5	5.8	0.24	0.060
ITBS Vocabulary Standard Score	200.8 (712)	196.9 (587)	198.0	199.0	1.0	0.05	0.462
TCAP Reading/LA Scale Score	507.2 (800)	508.5 (665)	509.2	506.6	-2.6	0.08	0.262
TCAP Mathematics Scale Score	508.9 (800)	510.8 (664)	507.1	509.9	2.8	0.09	0.571
TCAP Science Scale Score	193.9 (776)	194.0 (662)	194.0	193.5	-0.5	0.03	0.761
TCAP Social Studies Scale Score	193.1 (775)	193.5 (661)	195.6	191.5	-4.1	0.25	0.202

^a Numbers in parentheses are the numbers of students in each group having valid test scores from the baseline 2006 administrations and the spring 2007 administrations.

The estimated MCLA impacts in the sixth grade were about the same magnitude as they were for all grades, ranging between -4.1 and 5.8. The impacts tended to be more positive, i.e., in favor of the MCLA treatment students. The impact on the ITBS comprehension subtest and the TCAP Social Studies scores had associated effect sizes over 0.2, but neither were statistically significant ($p < 0.05$). Overall for sixth grade, there is no reason to reject the hypothesis that the average achievement of students in the treatment and control schools was the same at the end of Year 1. Enrollment in MCLA treatment schools did not have a significant impact on student achievement levels in reading or in the four core content areas in the sixth grade.

Similar analyses carried out for seventh- and eighth-grade students yielded non-significant estimated impacts with the exception of the impacts on ITBS Total Reading in the seventh grade and TCAP Reading/LA in the eighth grade. Both of these impacts favored students in control schools. The complete results of the multi-level analyses of the MCLA impact on these seven test scores for each grade can be found in Appendix J in Tables J15 – J35.

Immediate Impact of MCLA at Sixth Grade in Year 2

The immediate impact of MCLA on ITBS and TCAP test scores for sixth-grade students in Year 2 was determined using the same multi-level model (see Appendix G) for the 1,439 sixth-grade students in the Year 2 sample. The only differences were that the dependent variables were the 2008 ITBS and TCAP test scores and the respective test score covariate was from the 2007 administrations. The results of these analyses are summarized in Table 55. The complete results of the multi-level analyses of the Year 2 MCLA immediate impact on these seven test scores for sixth grade can be found in Appendix J in Tables J36 – J42.

Table 55: Immediate Year 2 Impact of MCLA on Spring 2008 Scores on Achievement Tests—Grade 6

Test Score	Unadjusted Means		Adjusted Means		Est. Impact	Effect Size ^a	Signif. Level
	Control	MCLA	Control	MCLA			
ITBS Total Reading Standard Score	197.6 (556) ^b	193.3 (469)	194.4	195.1	0.7	0.03	0.608
ITBS Comprehension Standard Score	196.8 (558)	192.2 (474)	192.8	194.5	1.8	0.07	0.280
ITBS Vocabulary Standard Score	198.8 (573)	194.3 (476)	195.8	196.4	0.7	0.03	0.668
TCAP Reading/LA Scale Score	513.6 (702)	512.5 (610)	510.3	514.7	4.4	0.13	0.250
TCAP Mathematics Scale Score	511.0 (705)	510.6 (610)	513.4	508.8	-4.6	0.14	0.478
TCAP Science Scale Score	195.3 (687)	191.7 (610)	195.2	191.9	-3.4	0.21	0.094
TCAP Social Studies Scale Score	192.4 (684)	191.3 (604)	191.9	190.8	-1.1	0.06	0.810

^a The method used to calculate effect size was Glass's Δ , the difference between treatment and control groups' adjusted mean test scores divided by the control group's test score standard deviation.

^b Numbers in parentheses are the numbers of students in each group having valid test scores from the baseline 2007 administrations and the spring 2008 administrations.

The estimated MCLA impacts in the sixth grade in Year 2 were of a similar size as those in Year 1. There were no significant impacts. Overall for sixth grade, there is no reason to reject the hypothesis that the average achievement of students in the treatment and control schools was

the same at the end of Year 2. MCLA participation did not have a significant impact on student achievement levels in reading or in the four core content areas in the sixth grade in Year 2. Thus, there was no variation in the impact of MCLA between Years 1 and 2; there were no impacts on any of the student outcome measures in either year.

Long-Term (Two-Year) Impact of MCLA at Grades 7 and 8 in Year 2

The long-term impact of schools participating in MCLA for two years was examined in two different ways. First, a two-level model similar to the model described in Appendix G was used to estimate the difference between students in treatment and control schools on spring 2008 ITBS and TCAP test scores, controlling for their 2006 baseline scores, along with other student-level covariates. Since the students in these analyses were the 2,335 “stayers” in the seventh and eighth grades in Year 2, only one dummy variable indicating which students were in eighth grade was employed to control for within-school differences attributable to the student’s grade level. The school-level covariates remained the same. Table 56 presents the results of these cross-sectional analyses of the two-year impact of MCLA.

Table 56: Long-Term (Two-Year) Impact of MCLA on Spring 2008 Scores on Achievement Tests

Test Score	Unadjusted Means		Adjusted Means		Est. Impact	Effect Size ^a	Signif. Level
	Control	MCLA	Control	MCLA			
ITBS Total Reading Standard Score	213.6 (974) ^b	211.6 (794)	213.0	211.0	-2.0	0.08	0.260
ITBS Comprehension Stan- dard Score	211.0 (982)	208.8 (797)	209.7	208.9	-0.8	0.03	0.743
ITBS Vocabulary Standard Score	215.7 (993)	213.9 (806)	216.0	213.7	-2.3	0.09	0.191
TCAP Reading/LA Scale Score	524.6 (1193)	522.9 (1010)	528.5	518.5	-10.2	0.37	0.046
TCAP Mathematics Scale Score	532.50 (1191)	523.2 (1010)	534.5	521.7	-12.8	0.41	0.029
TCAP Science Scale Score	109.8 (1161)	192.7 (1003)	192.3	191.4	-0.9	0.05	0.307
TCAP Social Studies Scale Score	193.7 (1158)	190.8 (997)	195.3	189.6	-5.7	0.44	0.025

^a The method used to calculate effect size was Glass’s Δ , the difference between treatment and control groups’ adjusted mean test scores divided by the control group’s test score standard deviation.

^b Numbers in parentheses are the numbers of students in each group having valid test scores from the baseline 2006 administrations and the spring 2008 administrations.

The estimated impacts of two years’ participation in MCLA on the “stayers” in the seventh and eighth grades in Year 2 were all in favor of the students in the control schools, and three were statistically significant ($p < 0.05$). Students in control schools achieved significantly higher

scores on the TCAP Reading/LA, Mathematics and Social Studies tests with effect sizes of 0.34, 0.40, and 0.43, respectively. When the analyses were restricted to eligible struggling readers, a similar pattern of results emerged but with one additional significant impact in favor of students in the control schools on the TCAP Reading/LA assessment.

Similar analyses were carried out separately for each of the two grades. In seventh grade, there were significant impacts favoring the control schools on all three ITBS measures as well as the TCAP Mathematics and Social Studies assessments. In eighth grade, there were only two significant impacts, one favoring the control schools (TCAP Mathematics) and one favoring the MCLA schools (ITBS Vocabulary). The complete results of the multi-level analyses of the MCLA impact on these seven test scores for the “stayers” in the seventh and eighth grades can be found in Appendix J in Tables J43 – J70.

The second method for examining long-term impacts of MCLA was to look for treatment and control differences in the average amount of growth in achievement from baseline to the end of Year 1 and from the end of Year 1 to the end of Year 2. This was accomplished by employing a three-level regression model. Level 1 represents a student’s test score as a function of the year in which it was measured: 2006, 2007, or 2008, or as the baseline score plus the growth in Year 1 plus the growth in Year 2. Level 2 represents the student’s baseline score as a function of the same student demographic covariates used in the previously described two-level analyses. Level 3 models the baseline score as a function of the school’s participation in MCLA and the same previously used school covariates. Also at level 3, the two growth coefficients are expressed as a function of the school’s participation in MCLA. (See Appendix K for a complete specification of this model.)

Table 57 presents the key results of these longitudinal analyses of the two-year impact of MCLA on growth in achievement for the 2,335 “stayers” in the seventh and eighth grades in Year 2. The table presents the average growth in test scores for students in MCLA treatment and control schools in Years 1 and 2 for each of the seven test scores. Underlined numbers indicate a significant ($p < 0.05$) difference between the amount of growth for students in treatment and control schools. The bold number indicates the greater growth in each pair.

Table 57: Comparison of Growth in Mean Test Scores for Students in Grades 7 and 8 in Year 2 in MCLA and Control Schools

Test Score	Year 1 Growth		Year 2 Growth	
	Control	MCLA	Control	MCLA
ITBS Total Reading Standard Score	2.05	1.93	9.74	10.37
ITBS Comprehension Standard Score	0.69	1.28	10.02	9.15
ITBS Vocabulary Standard Score	3.48	2.63	<u>9.18</u>	<u>11.18</u>
TCAP Reading/LA Scale Score	12.09	13.19	11.83	10.60
TCAP Mathematics Scale Score	<u>16.63</u>	<u>11.34</u>	<u>13.50</u>	<u>11.06</u>
TCAP Science Scale Score	<u>3.25</u>	<u>1.13</u>	<u>-1.81</u>	<u>0.35</u>
TCAP Social Studies Scale Score	-0.80	-1.98	<u>0.13</u>	<u>-1.22</u>

Growth in mean test scores is based on the 1048 students in MCLA schools and the 1,287 students in control schools from the Year 1 sample who were enrolled in an MSRP school from the same design group for a majority of instructional days during Year 2 and were enrolled for the spring 2008 ITBS administration (the “stayers”). Significantly different ($p < 0.05$) growth in mean test scores in Year 1 or Year 2 between students in MCLA and control schools are indicated by underlined numbers. The bold number indicates the greater growth in each pair.

The results of these growth analyses were mixed. The students in control schools demonstrated significantly ($p < 0.05$) greater growth in Year 1 on TCAP Mathematics and Science and in Year 2 on TCAP Mathematics and Social Studies. The students in MCLA schools had significantly greater growth in Year 2 on ITBS Vocabulary and TCAP Science. When these analyses were repeated for just the eligible struggling readers, there were no significant differences between the growth of students in treatment and control schools in either year. Similar analyses carried out separately for each of the two grades yielded several patterns of significant differences in the growth of students in treatment and control schools. In seventh grade, the students in the MCLA schools had significantly greater growth in Year 1 on ITBS Comprehension, while students in control schools demonstrated significantly greater growth in Year 2 on four of the seven measures: ITBS Total Reading and Comprehension and TCAP Mathematics and Social Studies (see Table 58).

Table 58: Comparison of Growth in Mean Test Scores for Students in Grade 7 in Year 2 in MCLA and Control Schools

Test Score	Year 1 Growth		Year 2 Growth	
	Control	MCLA	Control	MCLA
ITBS Total Reading Standard Score	1.07	2.50	<u>12.19</u>	<u>9.11</u>
ITBS Comprehension Standard Score	<u>-2.25</u>	<u>1.64</u>	<u>14.09</u>	<u>8.09</u>
ITBS Vocabulary Standard Score	4.51	3.26	10.04	9.98
TCAP Reading/LA Scale Score	11.46	13.83	7.65	5.27
TCAP Mathematics Scale Score	13.33	13.65	<u>16.13</u>	<u>8.13</u>
TCAP Science Scale Score	2.78	1.21	-1.36	-1.06
TCAP Social Studies Scale Score	-3.39	-2.27	<u>0.11</u>	<u>-4.46</u>

Growth in mean test scores is based on the 471 grade 7 students in MCLA schools and the 611 students in control schools from the Year 1 sample who were enrolled in an MSRP school from the same design group for a majority of instructional days during Year 2 and were enrolled for the spring 2008 ITBS administration (the “stayers”). Significantly different ($p < 0.05$) growth in mean test scores in Year 1 or Year 2 between students in MCLA and control schools are indicated by underlined numbers. The bold number indicates the greater growth in each pair.

In eighth grade, the students in control schools demonstrated significantly greater Year 1 growth in TCAP Mathematics, Science, and Social Studies, while students in MCLA schools grew significantly more in Year 2 on all three ITBS measures and on TCAP Science (see Table 59).

Table 59: Comparison of Growth in Mean Test Scores for Students in Grade 8 in Year 2 in MCLA and Control Schools

Test Score	Year 1 Growth		Year 2 Growth	
	Control	MCLA	Control	MCLA
ITBS Total Reading Standard Score	2.97	1.57	<u>7.50</u>	<u>11.53</u>
ITBS Comprehension Standard Score	3.49	1.07	<u>6.27</u>	<u>10.15</u>
ITBS Vocabulary Standard Score	2.44	2.21	<u>8.36</u>	<u>12.26</u>
TCAP Reading/LA Scale Score	12.89	12.43	15.78	14.72
TCAP Mathematics Scale Score	<u>19.91</u>	<u>9.59</u>	11.02	13.32
TCAP Science Scale Score	<u>3.53</u>	<u>0.93</u>	<u>-2.20</u>	<u>1.28</u>
TCAP Social Studies Scale Score	<u>1.66</u>	<u>-1.58</u>	0.10	1.17

Growth in mean test scores is based on the 577 grade 8 students in MCLA schools and the 676 students in control schools from the Year 1 sample who were enrolled in an MSRP school from the same design group for a majority of instructional days during Year 2 and were enrolled for the spring 2008 ITBS administration (the “stayers”). Significantly different ($p < 0.05$) growth in mean test scores in Year 1 or Year 2 between students in MCLA and control schools are indicated by underlined numbers. The bold number indicates the greater growth in each pair.

The complete results of the multi-level analyses of the MCLA impact on growth in these seven test scores for the “stayers” in the seventh and eighth grades can be found in Appendix L in Tables L1–L28.

Differential Attrition in Immediate and Long-Term MCLA Impact Analyses

The amount of attrition for the MCLA Year 1 immediate impact analyses was considerably smaller than it was for the *READ 180* impact analyses. For example, on the ITBS Total Reading score there was 23 percent attrition in the *READ 180* impact ITT sample, and on the TCAP scores attrition ranged from 5 to 10 percent (see Table C1 in Appendix C). However, for the MCLA impact sample, the attrition for ITBS Total Reading scores was only 14 percent and attrition on the TCAP scores ranged from 2 to 5 percent (see Table C2 in Appendix C).

Although the attrition is lower, there is still the potential for some bias in the estimated MCLA impacts due to differential attrition between the students in treatment and control schools, especially for the ITBS. This potential differential attrition was studied by comparing the average baseline 2006 test scores of the students who also had a spring 2007 score to the average of all students with baseline 2006 test scores, the difference being attributable to the attrition of stu-

dents. This comparison was done for students in both the treatment and control schools.²⁵ If this attrition effect were higher or lower in one group of schools, this differential attrition would have to be acknowledged as possibly biasing the estimated impact of MCLA.

The results of the study of differential attrition for the immediate impacts of MCLA may be found in Table C2 of Appendix C.²⁶ In summary, Table C2 shows that the effects of attrition in both treatment and control groups on baseline 2006 test scores did not exceed one standard or scale score point, and only one differential treatment effect was statistically significant ($p < 0.05$), on the TCAP Social Studies assessment, favoring the control students.

Conclusions

In the results of the analyses of immediate impacts in Year 1, there were no statistically significant differences between MCLA treatment and control students that were positive, i.e., favoring the MCLA treatment, and all of the estimated effect sizes were less than 0.2. The one statistically significant impact on TCAP Reading/LA scores was in favor of the students in the MCLA control schools. This outcome, however, was countered by an impact on ITBS comprehension scores that approached statistical significance that was in favor of the students in the MCLA treatment schools.

The results of the analyses of immediate impacts in Year 2 on sixth-grade students showed no significant impacts in favor of treatment or control schools. There was no difference between results for sixth grade students in Years 1 and 2. The results of the cross-sectional analyses of long-term impacts on students in the seventh and eighth grades in Year 2 showed seven significant differences, all but one in favor of the control schools. Longitudinal analyses of the long-term impact of MCLA on student growth revealed a mixed bag of results, with more significant impacts favoring control than treatment schools.

The lack of consistency of immediate impacts of MCLA participation on these student outcomes is not surprising. The logic model underlying the MCLA professional development model assumes that, first, teachers will improve their ability to provide literacy instruction in their classrooms and, second, this improvement will lead to improvements in student learning in the core content areas. Thus, despite the indications of a positive MCLA impact on teacher outcomes in Years 1 and 2 described earlier, to expect improvements in student achievement in the same school years may be overly optimistic.

Further Analyses

Exploratory analyses are planned that will study the relationship between MCLA impact and the amount of teacher participation in MCLA professional development in the four schools randomly assigned to the treatment group.

²⁵ The analytical method was a univariate ANOVA of baseline 2006 test scores, employing a 2x2 factorial design crossing the MCLA treatment/control condition with possession (yes/no) of a spring 2007 test score. The interaction of these two factors was tested for significance to determine whether or not there was a differential attrition effect.

²⁶ Attrition effects are reported only for one of the ITBS test scores—Total Reading—since the other two subtest scores are very highly correlated with the Total Reading score, and results would be expected to be very similar.

Summary of Year 2 RBS Data Collection Activities

Figure 10 summarizes the implementation and impact data collection activities conducted in Year 2. Information presented here was culled from various sources, including surveys; individual and focus group interviews; classroom observations; and reviews of *READ 180* documentation, coaching logs, and MCLA curricular resources.

Figure 10: Characteristics of Year 2 RBS Data Collection Methods

Data Collection Method and Topic	Date Conducted	Sample size*
Surveys		
Baseline characteristics and content knowledge—all content teachers	August 2007	N = 66 (teachers who had not previously completed a Year 1 survey)
Follow-up characteristics and content knowledge—all content teachers	May 2008	N = 169 (68.9%)**
MCLA teacher participants	January 2007	N = 48 (69.6%)
Interviews		
Striving Readers School Principals	September 2007	N = 8 (100%)
MCLA Treatment School Principals	May 2008	N = 4 (100%)
Literacy coaches	September 2007 May 2008	N = 6 (100%) N = 6 (100%)
MCLA instructors	Sept/October 2007 January 2008 May 2008	N = 4 (100%) N = 4 (100%) N = 4 (100%)
MCLA dropouts	May 2008	N = 7 (41.1%)
MCLA semester one focus group sessions	October 2007	N = 8 (62 teachers)
MCLA semester two focus group sessions	March 2008	N = 9 (53 teachers)
Observations		
Classrooms of control and treatment teachers	October 2007	N = 48
Classrooms of treatment teachers only (paired obs.)	May 2008	N = 10
<i>READ 180</i> classrooms— baseline	September 2008	N = 16 (84.2%)
<i>READ 180</i> classrooms— mid-year	February 2008	N = 15 (78.9%)
<i>READ 180</i> classrooms— follow-up	May 2008	N = 9 (47.3%)
MCLA evening course sessions	Periodic	N = 12
Student assessment		
Baseline ITBS	September 2007	N = 1,474
Follow-up ITBS	May 2008	N = 4,531

Data Collection Method and Topic	Date Conducted	Sample size*
Secondary Data		
MCLA attendance rosters, <i>READ 180</i> meeting attendance sheets	Year 2	All available data
Coaching calendar and log entries	Year 2	N = 6 coaches (100%)
TCAP	Spring 2008	N = 7,293 (all students, regardless of Year 2 school attended whose Year 1 school TCAP data were analyzed)

*Where possible, response rates are provided in parentheses.

**Calculated using MCS data file that lists all MSRP content and special education non-self-contained teachers (October 2007) where total N = 245 teachers is the denominator used to calculate percentages.

References

- Feldman and Feighan (2007). *Memphis Striving Readers Classroom Observation Protocol (MSR-COP)*. Philadelphia: RBS.
- Hall, G. E., and Hord, S.M. (2006). *Implementing Change: Patterns, Principles, and Potholes (Second Edition)*. Boston, MA: Allyn & Bacon.
- Hart, T. (2008, June 10). Memphis school board realizes it's time for change. *The Commercial Appeal*. Retrieved July 14, 2008, from <http://www.commercialappeal.com/news/2008/jun/10/guest-column-city-school-board-realizes-its-time/>
- MCS Quick Facts. (2008) Retrieved December 12, 2008 from http://www.mcsk12.net/facts_about_mcs.asp
- Rui, N. and Feldman, J. (2008). Inter-rater Reliability of a Classroom Observation Protocol: A Critical Appraisal. To be presented at the 2009 Annual Meeting of the American Education Association in San Diego, CA.
- Cooter, R., Potts, R., Feldman, J., Chadwick, K., Heeren, E., Perkins, H., Washington, R., Bryant, L., Harris, R., Allen, L., McCann, R., Hall, G. (2008). *Memphis Content Literacy Academy (MCLA) Innovation Configuration Map*. RBS: Philadelphia, PA.
- Potts, R., Perkins, J.H., Heeren, E., Harris, R., and Feldman, J. (2008, May). *Will that work for us? Interpreting research from the MSRP (MSRP)*. Paper presented at the International Reading Association Research Conference, Atlanta, GA.
- Scholastic (2005). *Leadership Implementation Guide: Supporting READ 180 in Your District*. Scholastic Inc.,: New York.
- Scholastic (2007). *READ 180 Enterprise Edition Research Protocol*. Scholastic Inc.,: New York.
- Urban Child Institute, The. (2008). *The state of children in Memphis and Shelby County*. Memphis, TN: Author. Retrieved July 21, 2008, from http://www.theurbanchildinstitute.org/data_book/processor.php