Cultivating Learning and
Positive Change

A Summary of the Treatment Group Findings from the Evaluation of Houghton Mifflin Harcourt Supplemental's *Literacy by Design* Program

October 1, 2008

magnolia consulting

Magnolia Consulting, LLC
3171 Courthouse Road
Louisa, Virginia 23093

phone 540.967.5540
fax 540.967.5541
info@magnoliaconsulting.org
www.magnoliaconsulting.org

EXECUTIVE SUMMARY

Helping children develop proficient reading skills is arguably the most critical task that educators face today. Being able to read fluently and effectively enables children to acquire knowledge in many content areas and is associated with academic achievement and future occupational success (Slavin, Karweit, & Madden, 1989). In accordance with the No Child Left Behind Act of 2001 (NCLB), which advocates for the use of educational methods that have been proven effective, Houghton Mifflin Harcourt Supplemental incorporated well-researched elements in its *Literacy by Design* program.

Literacy by Design is a comprehensive K–5 literacy program based on the gradual release of responsibility model (Pearson & Gallagher, 1983). *Literacy by Design* connects whole-class, small-group, and independent instruction through a common comprehension strategy. *Literacy by Design* covers 160 instructional days and is broken out into sixteen themes. The program provides direct instruction in phonics, vocabulary, and fluency, and it includes many different opportunities for students to practice skills in these areas. *Literacy by Design* connects instruction to content areas through texts linked to science and social studies themes that correlate to grade-level national standards.

Houghton Mifflin Harcourt Supplemental understands the importance of demonstrating the efficacy of its curriculum materials, and thus, contracted with Magnolia Consulting, LLC, an external, independent consulting firm specializing in educational research and evaluation, to conduct an efficacy study of its *Literacy by Design* curriculum materials.

STUDY PURPOSE AND FOCUS OF THIS REPORT

The purposes of the efficacy study were to evaluate teachers' perceptions and implementation of the *Literacy by Design* program and to assess the effectiveness of the materials in helping students attain critical reading skills.

This report focuses only on findings pertaining to a subset of the study sample—that is, users of *Literacy by Design*. Evaluators also developed a separate report that details findings pertaining to the full sample.

METHODOLOGY

The final treatment group analytical sample consisted of 11 treatment teachers, 3 site coordinators, and 199 third-grade students across three school districts in varying demographic regions. For the purposes of this study, the treatment teachers implemented the *Literacy by Design* program instead of their regular core reading programs during the 2007–2008 school year.

Evaluators utilized a variety of data collection techniques for treatment participants that focused on student and teacher characteristics, program implementation, and student learning. In particular, evaluators employed the following data collection mechanisms for treatment participants:

1) Fall, winter, and spring administrations of the *Gates McGinitie Reading Test, Fourth Edition®* (*GMRT-4)* student reading assessment;
2) fall and spring administration of the *Rigby Reading Evaluation and Diagnostic System*™ *(READS)* assessments;

3) online teacher survey (one administration in the fall and one administration in the spring);
4) spring classroom observations and interviews; and
5) teacher and student demographic information and student attendance.

FINDINGS PERTAINING TO TREATMENT GROUP

Key Questions:
How do teachers implement the* Literacy by Design *program? What are their perceptions of the program's quality and utility?

The findings indicate that treatment teachers implemented the *Literacy by Design* program appropriately in classrooms. They used it for 4–5 days per week; incorporated whole-group, small-group, and independent reading instruction; and used all of the core materials.

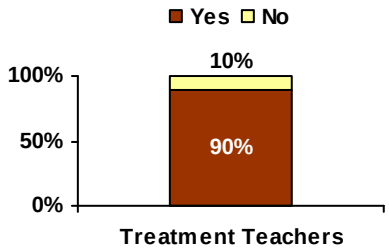
Key Finding:
Teachers implemented the *Literacy by Design* program with fidelity.

Findings from the online surveys, classroom observations, and interviews revealed that teachers generally liked the *Literacy by Design* program. They really appreciated the wealth of materials, as well as the program's structure, and they found the program enjoyable and engaging for their students. As a group, teachers generally thought the materials were effective at increasing student learning, and all teachers reported that *Literacy by Design* allowed them to meet most of students' needs. An overwhelming majority of the teachers indicated that they would like to continue using the program next year.

Teacher Quote:
I feel that individual needs are being met more adequately and I know that I am not missing skills or concepts to be taught. I like how it is interwoven with science and social studies.

Key Finding:
The majority of teachers perceived the program to be of high quality and utility and will use the program in the future.

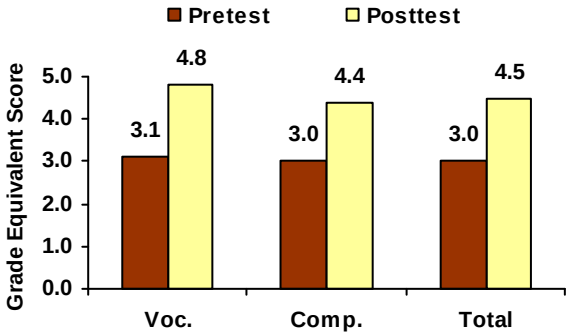
As a group, the treatment teachers seemed very pleased with *Literacy by Design*. Ninety percent of treatment teachers indicated that they would like to continue using the reading program next year.



Percentage of treatment teachers who would like to continue using *Literacy by Design*.

Key Question:
Do students in treatment groups demonstrate significant learning gains in reading?

Students who participated in the *Literacy by Design* program gained more than one grade level in their vocabulary and comprehension reading skills, on average, over course of the study. As a group, they scored at grade level on the *GMRT-4* at the beginning of the study, and they scored well above grade level by the end of the study.



Pretest and posttest grade equivalent scores for students exposed to *Literacy by Design*.

These students demonstrated significant large gains on the *GMRT* Vocabulary test (d = 1.70), Comprehension test (d = 1.14), and Total test (d = 1.69), translating to 45, 37, and 45 percentile points, respectively. *Literacy by Design* participants also demonstrated significant gains on the *READS*. In particular, treatment students demonstrated medium effect sizes (d = 0.45) on the *READS* Reading Comprehension and Vocabulary in Context (d =0.46) subtests, translating to gains of 17 percentile points. Treatment students demonstrated large effect sizes (d = 0.61) on the *READS* Sounds-Letters Vowels and Word Part Clues (d = 0.66) subtests, translating to gains of 23 and 24 percentile points, respectively. Pretest/posttest differences on the *READS* Sounds-Letters Consonants subtest were not significant, and the effect sizes were small d = –0.18), translating to decreases of only 7 percentile points.

Teacher Quote:
The lower students are improving, and they have more confidence in what they are doing. They can figure out what the story is about instead of being frustrated.

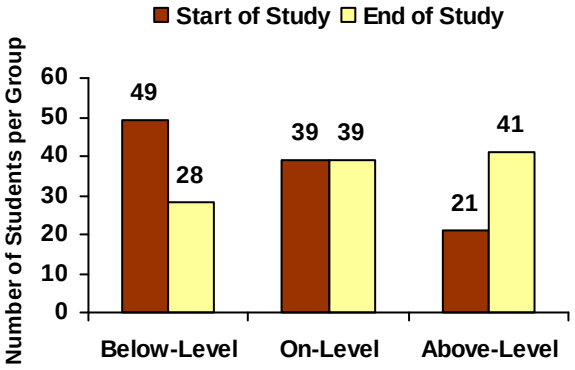
Key Finding:
Treatment students demonstrated significant gains (corresponding to medium and large effect sizes) on the *GMRT-4* and *READS* over the course of the study.

In addition to developing specific reading skills, the study's findings suggest that students who participated in the *Literacy by Design* program became better able to receive instruction and read independently at higher levels over the course of the school year. Taken together, these results suggest that the *Literacy by Design* program is effective at significantly increasing student learning in reading.

Key Finding:
***Literacy by Design* is effective at significantly increasing student learning in reading.**

Analyses examining the effectiveness of *Literacy by Design* with students of varying reading abilities revealed that within each ability level group (i.e., below-level, on-level, and advanced-level), students demonstrated significant gains with large effect sizes on the *GMRT-4* Vocabulary, Comprehension, and Total tests. Further analyses revealed that there were no significant differences in learning gains by ability group. This important finding suggests that the *Literacy by Design* program works equally well for students, regardless of their ability level. Thus, *Literacy by Design* is appropriate for use in classrooms where students differ in their reading abilities and literacy skills. Additionally, subsequent analyses revealed that many students who participated in *Literacy by Design* moved to higher ability groups over the course of the school year, which confirms earlier findings that the program is effective at increasing students' reading skills.

Key Finding:
***Literacy by Design* works equally well for students of varying ability levels, making it appropriate for use in classrooms in which students have diverse literacy skills.**



Numbers of treatment students in each ability group at pretest and posttest

In conclusion, teachers who participated in the *Literacy by Design* program found it effective and engaging. They appreciated the structure of the program, as well as the comprehensiveness of the materials. The findings of this study indicated that during the first year of implementation, *Literacy by Design* was successful in significantly improving children's reading skills in the areas of vocabulary, comprehension, and fluency, as well as helping children become better able to read higher levels of text independently.

ACKNOWLEDGEMENTS

This evaluation study represents a shared effort among Magnolia Consulting, Houghton Mifflin Harcourt Supplemental, and study participants across three school districts. Magnolia Consulting evaluators are appreciative of everyone whose contributions and support made this work possible. We especially want to offer our sincerest appreciation to all of the study participants, including teachers, students, administrators, literacy support staff, and site coordinators. Each of these individuals graciously responded to Magnolia's requests, shared experiences, and provided valuable feedback throughout the course of the project. We also gratefully acknowledge staff at Houghton Mifflin Harcourt Supplemental for their ongoing support and understanding of the important nuances of applied research. Finally, we would like to convey sincere gratitude to the support and contributions provided by Kristine Chadwick and Chandra O'Connor for their contributions to site management, as well as Arianne Welker and Mary Styers of Magnolia Consulting.

The authors,

Lisa C. Shannon, Ph.D.
Stephanie Baird Wilkerson, Ph.D.

TABLE OF CONTENTS

Executive Summary ..... *i*

Acknowledgements ..... *v*

Introduction 1

Focus of this Report 2

Methodology 2
- Measures 3
- Procedures 3

Site Overview 5
- Treatment Participants 5

Description of *Literacy by Design* 7
- Theoretical Foundation 7
- Program Materials and Components 8

Literacy by Design Program Implementation and Teacher Perceptions 8
- *Literacy by Design* Implementation 9

Student Performance Results 21
- *Literacy by Design* Impacts on Student Learning 21

Conclusion 32

References 34

Appendix A: Study Design and Methodology 35

Appendix B: Data Collection Activities 38

Appendix C: Assessment Information 39

Appendix D: Implementation Guidelines 42

Appendix E: Site Characteristics 45

Appendix F: Treatment Student Characteristics 47

Appendix G: Supporting Tables for Program Implementation 48

Appendix H: Supporting Tables for Student Performance Results 49

List of Tables

Table 1: Timeline of Treatment Group Data Collection Activities 4
Table 2: Mean *GMRT* Gains for Treatment Students 26
Table 3: Paired Samples t-Tests for *GRMT-4* Gains by Ability Group 29
Table 4: Mean *GMRT* Gains for Treatment Students 31
Table A1: Data Analysis Plan Summary for Treatment Group Analysis 36
Table C1: Reliability Information for the *GMRT-4* 39
Table C2: Reliability Information for the *READS* 40
Table E1: Site Characteristics by District 45
Table F1: Treatment Student Characteristics 47
Table G1: Treatment Teacher Average use of Whole-class Instruction, Small-group instruction and Professional Development Materials 48
Table H1: *GMRT-4* Gains by Treatment Student Characteristics 49
Table H2: *READS* Gains by Treatment Student Characteristics 50
Table H3: Treatment Group Paired Samples t-Tests for *READS* Gains by Ability Group 52
Table H4: Treatment Group *GMRT-4* Gains by Site 53
Table H5: Treatment Group *READS* Gains by Site 54

List of Figures

Figure 1: Planning and preparation time by survey administration 9
Figure 2: Days per week spent on each type of instruction 10
Figure 3: Minutes per day spent on each type of instruction 10
Figure 4: Percentage of teachers who report assessing their students at various times 12
Figure 5: Percentage of teachers taking specific actions because of assessments 12
Figure 6: Percentage of teachers who reported using supplemental materials 13
Figure 7: Teacher perceptions regarding ease of program implementation 14
Figure 8: Teachers' perceptions regarding utility of training 14
Figure 9: Teachers' perceptions regarding the amount of material offered by *Literacy by Design* 15
Figure 10: Teachers' perceptions regarding the pacing of the material offered by *Literacy by Design* 15
Figure 11: Adequacy of *Literacy by Design* in meeting the needs of students at various levels 16
Figure 12: Program effectiveness in increasing student learning across survey administrations 17
Figure 13: Teachers' perceptions of student engagement across survey administration 17
Figure 14: *GMRT-4* pre/post grade equivalent scores 22
Figure 15: Pretest *READS* Instructional Reading Level 23
Figure 16: Posttest *READS* Instructional Reading Level 23
Figure 17: Pretest *READS* Independent Reading Level 23
Figure 18: Posttest *READS* Independent Reading Level 23
Figure 19: Pretest *READS* Proficiency Level 24
Figure 20: Posttest *READS* Proficiency Level 25
Figure 21: *GMRT-4* pre-, mid-, and posttest extended scale scores 25
Figure 22: *READS* pretest and posttest raw scores 25
Figure 23: Pre/post changes in numbers of students in each ability group 30

INTRODUCTION

Helping children develop proficient reading skills is arguably the most critical task that educators face today. Being able to read fluently and effectively enables children to acquire knowledge in many content areas and is associated with academic achievement and future occupational success (Slavin, Karweit & Madden, 1989). Increasingly, educators are appreciating the notion that how teachers provide instruction is just as important as the content they use for instruction and that utilizing a comprehensive, balanced approach to literacy, which combines whole-class and small-group instruction is a key element of effective instruction (NICHD, 2001; Neuman & Dickinson, 2001). Armbruster, Lehr & Osborne (2001) highlight the importance of utilizing direct and connected comprehension, phonics, phonemic awareness, vocabulary, and fluency instruction, and a number of studies suggest that differentiated and interactive instruction are key elements to effective reading instruction (e.g., see Tyner, 2004; Mathes, Torgesen, Clancy-Menchetti, Santi, Nicholas, Robinson, et al., 2003; Lou, Abrami & Spence, 2000).

In accordance with the No Child Left Behind Act of 2001 (NCLB), which advocates for the use of educational methods that have been proven effective, Houghton Mifflin Harcourt Supplemental (HMHS) incorporated well-researched elements in its *Literacy by Design* program. *Literacy by Design* is a comprehensive K–5 literacy program based on the gradual release of responsibility model (Pearson & Gallagher, 1983). This model posits that teachers progressively decrease their responsibility for the student's learning as the learner becomes more proficient. In *Literacy by Design,* this gradual release of responsibility occurs through Modeled Reading (in which the teachers shoulder most of the responsibility for students' learning), Shared Reading, Interactive Reading, Small Group Strategic Reading, and Independent Reading (in which the student shoulders most of the responsibility for his or her own learning) (Harcourt Achieve, 2007). *Literacy by Design* uses explicit and connected comprehension instruction to tie together whole class, small-group, and independent reading, and the program incorporates differentiated instruction via leveled readers, small group instruction, and specific tools designed to facilitate differentiated instruction, as well as tools to facilitate instruction with English language learners (ELL). *Literacy by Design* provides direct instruction in phonics, vocabulary, and fluency, as well as many different opportunities for students to practice skills in these areas. The program connects instruction to content areas through texts linked to science and social studies themes that correlate to grade-level national standards, and the variety of program materials seek to increase student motivation in reading. *Literacy by Design* offers tools and guidance for ongoing formative and summative assessment, which is critical to informing instruction.

Houghton Mifflin Harcourt Supplemental understands the importance of demonstrating the efficacy of its curriculum materials, and, thus, contracted with Magnolia Consulting, LLC, an external, independent consulting firm specializing in educational research and evaluation, to conduct an efficacy study of its *Literacy by Design* curriculum materials. Magnolia Consulting conducted this study for Houghton Mifflin Harcourt Supplemental during the 2007–2008 school year. This report presents a description of *Literacy by Design*, an overview of teachers' perceptions of the program, a description of treatment students' performance results, and a discussion of the findings pertaining to

teachers and students who used or who were exposed to the *Literacy by Design* program during the study.[1]

FOCUS OF THIS REPORT

The purpose of the efficacy study was to evaluate teachers' perceptions and implementation of the *Literacy by Design* program and to assess the effectiveness of the materials in helping students attain critical reading skills. Evaluators conducted the study with third-grade participants across three school districts in varying geographic regions. This report details information pertaining only to the treatment group, which was comprised of teachers who implemented *Literacy by Design* in their classrooms and students who participated in the program throughout the course of the study.[2] This report focuses on the following overarching evaluation questions pertaining only to the study's treatment group:

1. How do teachers implement the *Literacy by Design* program?
2. What are teachers' perceptions of the *Literacy by Design* program?
3. Does the *Literacy by Design* program have a significant impact on student learning gains in reading over time?
4. How does the *Literacy by Design* program serve various subgroups of students?

METHODOLOGY

While conducting this study, evaluators collected both teacher and student data. Appendix B provides an overview of data collection activities for the efficacy study. Data collection focused on student and teacher characteristics, program implementation, and student learning. In particular, evaluators employed the following data collection mechanisms for the treatment participants in this study:

- fall, winter, and spring administrations of the *Gates McGinitie Reading Test, Fourth Edition® (GMRT-4)* student reading assessment for treatment students
- fall and spring administration of the *Rigby Reading Evaluation and Diagnostic System™ (READS)* assessment for treatment students
- two administrations of an online treatment teacher survey
- spring classroom observations and interviews
- teacher and student demographic information and student attendance

[1] This report represents findings pertaining to the treatment group of teachers and students who participated in a randomized controlled trial Magnolia Consulting conducted during the 2007–2008 school. Magnolia Consulting reported the findings from the RCT separately.
[2] Appendix A provides additional information about the study design and data analysis plan.

MEASURES

Evaluators used a variety of student and teacher measures in the conduct of this study. This portion of the report provides an overview of these measures.

Student Assessments

The *GMRT-4* is a group-administered assessment that gauges decoding skills, vocabulary, reading comprehension, and word knowledge. The Level 3 subtest, which is appropriate for third-grade students, yields Vocabulary, Comprehension, and Total test scores. Appendix C contains additional information about the *GMRT-4*.

The *READS* is a group-administered assessment that gauges reading comprehension, phonemic awareness and phonics, vocabulary, word part clues, skimming and scanning text for information, and fluency. The test yields *Instructional* and *Independent Reading Levels* that indicate the level of text a student can read and comprehend. The *READS Diagnostic Test for Grade 3* also yields separate scores for Reading Comprehension, Sounds-Letters: Consonants, Sounds-Letters: Vowels, Vocabulary in Context, and Word Part Clues. Appendix C contains additional information about the *READS*.

Teacher Measures

In the fall and spring, treatment teachers completed an online survey that asked questions about program implementation and provided teachers with an opportunity to provide feedback about *Literacy by Design.* Appendix C contains additional information about the online surveys.

Evaluators conducted site visits in the spring with all treatment teachers. During these visits, evaluators observed and interviewed participating teachers. These site visits enabled evaluators to triangulate other data sources and develop a deeper understanding of program implementation. Appendix C contains information about the observation protocols and interview protocols used during the site visits.

PROCEDURES

This portion of the report provides an overview of the study's procedures. Evaluators adhered to explicit procedures to execute the study efficiently and effectively.

Site Selection

Evaluators and Houghton Mifflin Harcourt Supplemental developed preferred criteria for potential sites, which included the following: (1) interest in implementing the *Literacy by Design* program in grade three, (2) high attendance and low mobility rates, (3) use of a competing basal or comprehensive reading curriculum, (4) low ESL population, (5) adequate reading block available to dedicate to *Literacy by Design* instruction, (6) geographic distribution, and (7) comfort with teacher-level random assignment. Because differentiated instruction is an integral component of the *Literacy by Design* program, evaluators also focused on selecting sites with students demonstrating various levels of reading abilities as opposed to schools where a majority of students perform above or below grade level.

Evaluators engaged in multiple efforts to find the largest number of potential sites for the study. Specifically, Magnolia Consulting pulled from its extensive database of over 1,000 district contacts and identified those with an interest in participating in a study during the 2007–2008 school year. Magnolia Consulting also advertised the study through multiple formats, including e-mail and its website. Once evaluators identified potential sites, evaluators contacted sites about the study and sent them a study description and application. Evaluators then followed up with interested sites via phone, e-mail, and facsimile.

Treatment Group Data Collection Timeframe

Table 1 displays the data collection timeframe evaluators followed for this study. First, Magnolia Consulting conducted a study orientation in conjunction with Houghton Mifflin Harcourt Supplemental's *Literacy by Design* program training. Within a few days of participating in this training, teachers administered the pretest assessments to their students. Then, teachers began implementing the *Literacy by Design* program in their classrooms. In October, evaluators administered the first teacher online survey. In January, teachers administered the mid-study *GMRT-4* student assessment. Evaluators conducted site visits, which consisted of classroom observations and interviews with all teachers. In April, evaluators administered the second teacher online survey. In May, teachers administered the posttest assessments (i.e., treatment teachers administered the *GMRT-4* and *READS*). The study period officially ended in June. The study period did not coincide with disruptive events at Site 1, Site 2, or Site 3.

Table 1.
Timeline of Treatment Group Data Collection Activities

TASK AND ACTIVITY	August	September	October	November	December	January	February	March	April	May	June
Training, study orientation, study begins	Site 3	Sites 1 & 2									
Administration of pretest student measures	→	→									
Administration of treatment teacher online survey			X								
Administration of midyear *GMRT-4*						X					
Classroom observations and interviews								X			
Administration of treatment teacher online survey									X		
Administration of posttest student assessments										X	
End of study											X

Implementation Fidelity

Magnolia Consulting and Houghton Mifflin Harcourt Supplemental took several steps to ensure that treatment teachers implemented the *Literacy by Design* program appropriately during the study period. First, Houghton Mifflin Harcourt Supplemental provided program training to all participating teachers prior to the onset of the study. This training addressed the program components, materials, and implementation, and it provided teachers with an opportunity to ask questions. During the training, evaluators also provided teachers with implementation guidelines, which specified exactly how to implement the program with fidelity. Appendix D contains these specific implementation guidelines. Houghton Mifflin Harcourt Supplemental also conducted a follow-up training approximately six-to-eight weeks after program implementation began so that teachers could ask questions and gain clarity in any implementation issues they were experiencing. Throughout the study period, evaluators communicated with teachers and site coordinators via e-mail and phone to ensure that they were implementing the program appropriately and to respond to any questions or issues that arose. Evaluators also monitored teachers' program implementation through two online teacher surveys, which measured teachers' use of the materials and program components. Finally, evaluators conducted a site visit with all treatment teachers, which gave evaluators an opportunity to observe teachers implementing the program in their classrooms and conduct a follow-up interview with teachers. These methods not only ensured that teachers implemented the *Literacy by Design* program appropriately, but also provided teachers with an opportunity to ask questions and provide feedback about the *Literacy by Design* program.

SITE OVERVIEW

At the request of Houghton Mifflin Harcourt Supplemental, Magnolia Consulting recruited at least one school from each of three school districts located in varying geographic regions of the country. Recruitment efforts yielded four participating schools across three school districts for a total initial sample size of 23 third-grade teachers and 440 third-grade students. At the time of the final report, there were 417 third-grade student participants. Table E1, located in Appendix E, indicates that the three school districts recruited for the study—Site 1, Site 2, and Site 3—are diverse regarding their geographic location, size, and economic status. The sites have varied in their past performance on statewide assessments. Site 1 has two participating schools and Sites 2 and 3 each have one participating school.

TREATMENT PARTICIPANTS

The final treatment analytical sample for the study included 199 students and 11 teachers who implemented the *Literacy by Design* program in their classrooms. Additionally, each district identified a district-level site coordinator and a school-level coordinator. The district and school coordinators were primary contacts for study-related issues, and their responsibilities included distributing materials, ensuring that teachers administered assessments and implemented the program, and assisting with site visit scheduling.

Treatment Teacher Participants

The treatment teachers contributed to the study by implementing the *Literacy by Design* program in their classrooms, administering student assessments (including the *READS* and *GMRT-4*), completing two online surveys, and participating in spring observations and interviews. As an incentive for participation, teachers received a $300 stipend. Participating teachers also received free curriculum materials and training. All teachers and coordinators completed an informed consent form for the study.

Of the eleven treatment teachers, five held a bachelor's degree and five held a master's degree.[3] The teachers had been teaching for a range of one to thirty-eight years (χ = 12.55). They had taught at their current schools for a range of one to twenty-seven years (χ = 9.00). The number of students in the final analytical sample that worked with each of these eleven teachers ranged from 12 to 24.

Treatment Student Participants

As indicated previously, the final treatment analytical sample for this study consisted of 199 students in the third grade. This section presents a description of the students' demographic characteristics.

Demographics

Approximately one-half of the treatment students (52%) were male, and one-half (48%) were female. In the treatment group, approximately 2% of students were African-American, 7% were Hispanic, 87% were Caucasian, and 5% were categorized as either multi-racial, Asian, American Indian, or other. Of these students, 41% qualified for free or reduced-price lunch. Of the treatment sample, 7% included special education students, and 6% of the students were categorized by their district as limited English proficient (LEP). Of treatment students, 2% were categorized as Section 504,[4] and no participants were classified as migrant students. Table F1 in Appendix F displays these data.

Attrition

The initial treatment group study sample included 205 *Literacy by Design* participants. The final sample for analysis included 199 treatment participants. The overall attrition rate was 3%. Out of the six dropped treatment participants, two moved during the course of the study, three did not have complete pre-/post-GMRT or READS data, and one student attended less than 80% of classes.

[3] One teacher did not indicate whether she held a bachelor's degree or master's degree.

[4] A student categorized as Section 504 is one with an impairment that may include any disability, long-term illness, or various disorder (e.g. ADHD, diabetes, epilepsy, allergies) that *substantially* reduces or lessens a student's ability to access learning in the educational setting. These students may receive test accommodations and modifications. A physical or mental impairment does not constitute a disability for purposes of Section 504.

DESCRIPTION OF *LITERACY BY DESIGN*

Houghton Mifflin Harcourt Supplemental's *Literacy by Design* is a complete literacy program that connects whole-class, small-group, and independent instruction through a common comprehension strategy. The K–5 program links literacy skills to standards-based science and social studies content through the use of fiction and nonfiction texts and themes. Instruction focuses on listening, speaking, reading, and writing within content-specific themes, which simultaneously builds reading and content-area competence among students.

THEORETICAL FOUNDATION

A number of leading authors in the fields of reading and writing designed *Literacy by Design* to be a thorough and comprehensive literacy program that engages children, motivates them to learn, and exceeds national curriculum mandates to help students achieve adequate yearly progress. The program targets comprehension, vocabulary, fluency, phonics and phonemic awareness, and writing, and it also provides support materials to facilitate instruction for English language learners (ELL) (Harcourt Achieve, Inc. 2008).

Literacy by Design is grounded in research around key areas of literacy development and specific techniques. First, the program utilizes explicit comprehension instruction that links whole-class, small-group, and independent reading via common comprehension strategies, including the following: making connections, determining importance, inferring, using fix-up strategies, synthesizing, creating sensory and emotional images, asking questions, and monitoring understanding. The program also incorporates comprehensive and explicit writing instruction, covering all four essential elements of writing instruction: writing forms, organizational patterns, writing traits, and writing process. Furthermore, reading and writing instruction are integrated so that students learn to write as readers and read as writers, which will help students understand text and be better able to express themselves through writing (Harcourt Achieve, 2007).

Literacy by Design is premised on the gradual release of responsibility model. This model facilitates student independence in reading and writing by gradually reducing teacher support as students become more competent. In other words, teachers provide an appropriate amount of support and guidance and give students opportunities to practice and build confidence so they can eventually become independent readers and writers. *Literacy by Design* incorporates this model for reading through a number of techniques, including modeled reading, shared reading, interactive reading, small-group strategic reading, and independent reading (Harcourt Achieve, 2007).

In response to research showing the effectiveness of differentiated and interactive instruction, *Literacy by Design's* authors incorporated a very specific and easy to follow leveling system into the program (Harcourt Achieve, 2007). This leveling system allows teachers to provide differentiated instruction and enables students of all abilities to be successful. The program also incorporates many opportunities for interactive instruction in the form of small-group practice, partnering, interactive reading techniques, and other interactive practices. These techniques get students actively involved in the learning process and keep them engaged in their lessons.

Literacy by Design covers key literacy components by offering explicit and systematic instruction and support materials for phonics and phonemic awareness, vocabulary, and fluency development. The program also connects instruction to various content areas by focusing instruction on listening, speaking, reading, and writing in the context of science and social studies themes. The materials include fiction and nonfiction selections, and the program builds academic vocabulary to facilitate students' success across subjects. The content area stories and themes also serve to motivate students and keep them interested in the material (Harcourt Achieve, 2007).

Finally, in the *Literacy by Design* program, assessment drives instruction. Therefore, teachers have ample resources available to them to facilitate ongoing formative and summative assessment. Teachers can use the results of assessments to tailor their instruction to meet the needs of their students best, as well as to monitor students' progress (Harcourt Achieve, 2007).

PROGRAM MATERIALS AND COMPONENTS

Literacy by Design covers 160 instructional days. The program is broken out into 16 themes, each of which takes approximately two weeks to cover. The Teacher's Guide recommends that teachers spend approximately 90 minutes implementing the reading component of the program. If teachers implement the writing component, they should spend approximately 60 minutes doing so. Four levels of readers are available for Kindergarten, and eight levels of readers are available for each of Grades 1–5. The assessment components of *Literacy by Design* are tied closely to instruction and they allow for diagnostic assessment and ongoing progress monitoring.

The program is comprised of numerous resources to facilitate the flow of instruction. Specific program resources include a comprehensive Teacher's Guide, comprehension and writing bridge cards, differentiated instruction teacher's guides, student sourcebooks, leveled readers, reading and writing transparencies, charts, assessment guides, and other supporting materials.

LITERACY BY DESIGN PROGRAM IMPLEMENTATION AND TEACHER PERCEPTIONS

As indicated previously, evaluators provided teachers with implementation guidelines at the onset of the study. These guidelines asked teachers to use the program for a minimum of 90 minutes per day, five days a week, for 32 weeks. Evaluators gauged teachers' program implementation via two online surveys (one in the fall and one in the spring),[5] one site visit consisting of classroom observations and interviews, and via regular communications with site coordinators. This section of the report presents the findings pertaining to teachers' implementation of the *Literacy by Design* program, as well as teacher's perceptions of the program. Evaluators assessed teachers' program implementation based on the following: 1) program planning, preparation, and rate of implementation and 2) use of materials and type of instruction. Evaluators assessed teachers' perceptions of the program according to: 1) ease of implementation and utility of training, 2) amount of materials and pacing of instruction, and 3) perceptions of student learning and engagement.

[5] All 11 treatment teachers completed the fall online survey, and 10 treatment teachers completed the spring survey.

LITERACY BY DESIGN IMPLEMENTATION

This part of the report provides findings regarding teachers' implementation of the *Literacy by Design* program. It includes information regarding how they implemented the program, as well as their perceptions of the program.

Key Question:
How do teachers implement the* Literacy by Design *program?

Planning, Preparation, and Rate of Implementation

This section of the report provides a description of treatment teachers' planning time, preparation time, and rate of implementation of the *Literacy by Design* program. Evaluators measured these components via online surveys, teacher interviews, and classroom observations.

Planning and Preparation

As expected, teachers generally reported spending more time planning and preparing in the fall than in the spring (see Figure 1). On the fall survey, 63% of teachers reported spending over an hour per week planning and preparing, but on the spring survey, only 20% of teachers reported spending over an hour per week planning and preparing. This finding was not surprising, as teachers generally require less planning and preparation time as they gain comfort and familiarity with a program.

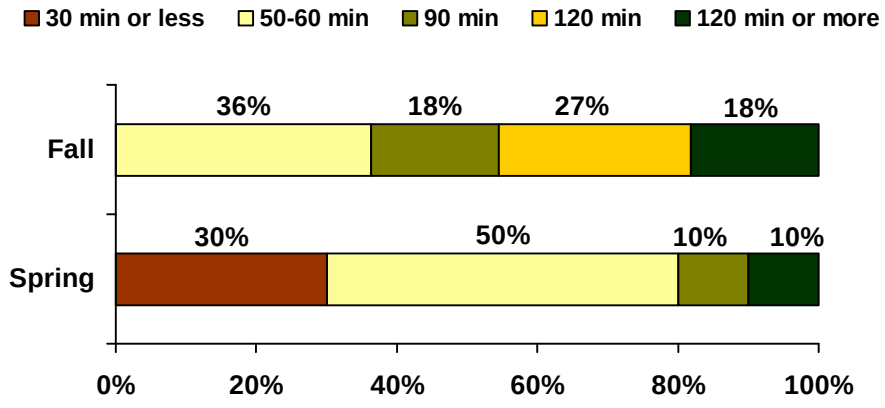


Figure 1. Planning and preparation time by survey administration.

Rate of Program Implementation

The implementation guidelines indicated that teachers should implement *Literacy by Design* for approximately 90 minutes per day (split between whole class instruction, small group instruction, and independent reading), five days per week. Findings indicate that overall, implementation across treatment teachers was quite high. On average, teachers reported using the program five days a week at the fall (χ = 4.91) and spring (χ = 5.00) surveys. On the fall survey, all teachers except for one reported using the program an average of five days per week, and on the spring survey, all teachers reported using the program an average of five days per week.

Regarding types of instruction, most teachers (76%) reported using whole-class instruction for modeled, shared, and interactive reading an average of five days each week. Similarly, the majority of teachers (67%) used small-group instruction for differentiated reading an average of five days per week. All teachers also reported that students engaged in independent reading an average of five days per week (see Figure 2).

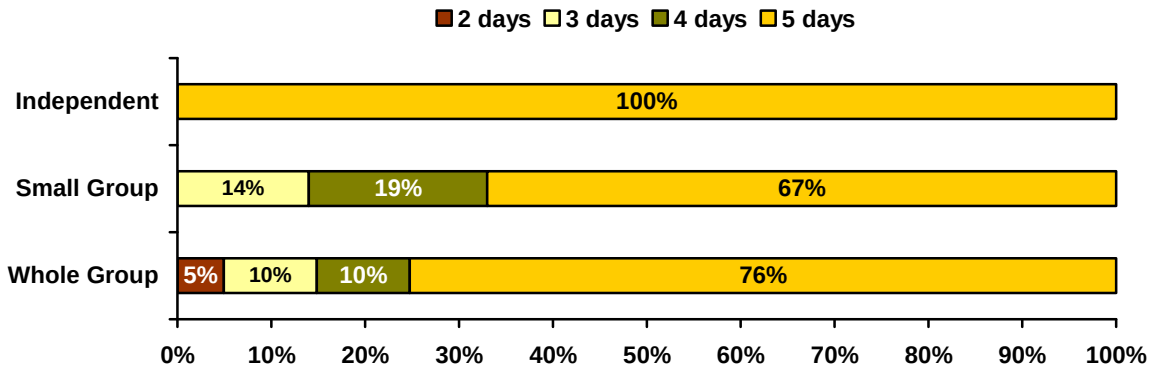


Figure 2. Days per week spent on each type of instruction.

Teachers varied in the duration in which they taught whole-group instruction, spending anywhere from 15–25 minutes per day to more than 55 minutes per day. However, the majority of teachers (90%) spent 30–55 minutes on small-group instruction and 15–25 minutes on independent reading (76%) (see Figure 3).

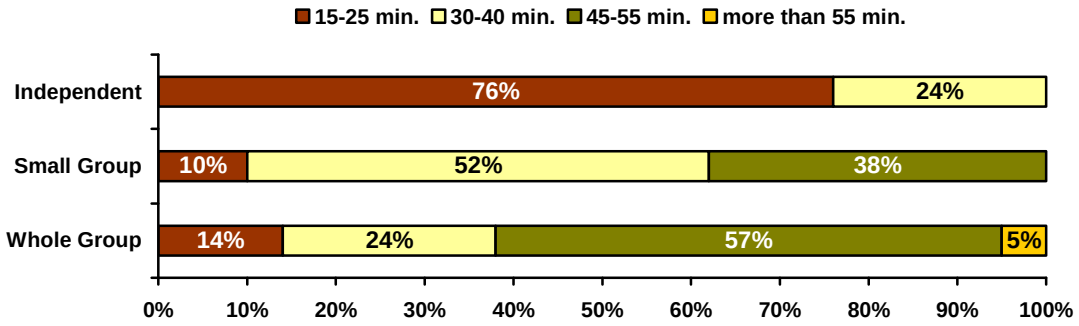


Figure 3. Minutes per day spent on each type of instruction.

These findings indicate that, as a whole, teachers implemented whole-group, small-group, and independent instruction for an appropriate number of minutes per day and for an appropriate number of days per week.

Materials Use

This section of the report provides an overview of treatment teachers' use of program materials. Evaluators measured teachers' materials use via online surveys, teacher interviews, and classroom observations.

Use of Core Program Materials

The *Literacy by Design* program is a comprehensive literacy program, and, as such, its materials are intended to serve as the core materials for reading instruction. The program includes materials for whole-class instruction, small-group instruction, as well professional development materials. Findings from surveys, interviews, and observations indicate that teachers generally used all of the core program materials. As expected, across both survey administrations, teachers most frequently used the *Comprehensive Teacher's Guide* and *Sourcebook* during whole-class instruction. Within the context of small-group instruction, teachers frequently used the *Small Group Reading Instruction Teacher's Guide* and *Leveled Readers*. Only one teacher reported accessing the professional development components (specifically, the online implementation training). The remaining teachers did not report utilizing these components. Table G1 in Appendix G displays additional information about core program materials use.

Assessment Material Use

As a group, teachers used all of the main *Literacy by Design* assessment materials. Teachers generally utilized the *Assessment Guide* a little over once per week, and, as expected, they used the *Benchmark Book Evaluation Guide, Benchmark Books,* and *READS* materials less frequently. Most teachers indicated that they assessed students at the beginning of the year, weekly, and at the end of each lesson/unit. Many teachers also indicated that they assessed their students quarterly (see Figure 4). Two teachers also indicated that they utilized informal assessments daily through observation of students, one teacher mentioned using standardized tests, and another teacher cited using assessment when students had completed their level. Table G1 in Appendix G displays additional information about Assessment materials use.

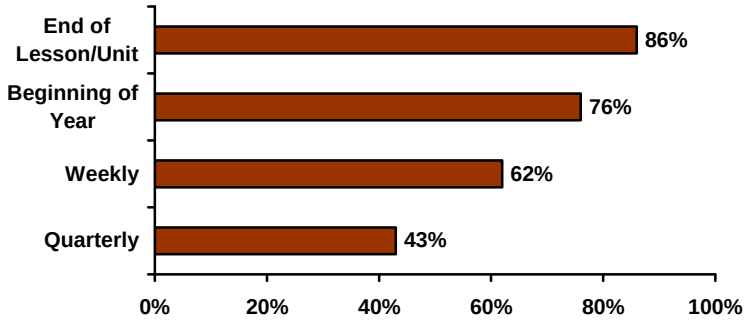


Figure 4. Percentage of teachers who report assessing their students at various times.

Teachers reported taking a number of actions resulting from their use of assessments. Figure 5 illustrates that nearly all teachers reported using the results of assessment practices to group students and to re-teach aspects of a lesson, and a majority of teachers indicated that they used assessment results to modify instruction. Clearly, *Literacy by Design* teachers used their assessments to make informed decisions about their instruction.

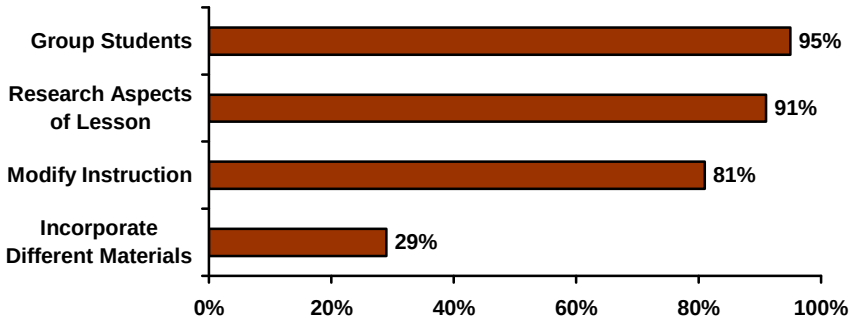


Figure 5. Percentage of teachers taking specific actions because of assessments.

Support and Technology Material Use

As mentioned previously, *Literacy by Design* includes a variety of support and technology materials. With regard to support materials, teachers most frequently used the skills master and independent reading poster across survey administrations. In contrast, teachers reported very little to no use of the technology materials (see Table G1 in Appendix G). Some teachers expressed an interest in using the technology materials but indicated that they were unable to access the online system. Evaluators worked with Houghton Mifflin Harcourt Supplemental to solve the technical problems, but they were not able to resolve the issues.

Supplemental Material Use

Participating teachers were required to use the *Literacy by Design* program as their core reading program, and evaluators asked teachers not to supplement the program unless it was necessary to meet students' needs. Figure 6 demonstrates that, in the fall, most teachers reported that they did not supplement the program, but by the spring, 60% (n = 6) of the teachers reported supplementing sometimes or often. Specifically, 50% of the teachers reported supplementing sometimes, and 10% reported supplementing the program often. Teachers who used supplementary materials reported adding additional materials for comprehension, word study, phonics, fluency, skills, grammar, and additional grades. In the spring survey, several teachers indicated that although they did not supplement the program, they would like to do so in the future. In particular, one teacher indicated a desire to supplement the activity pages to facilitate comprehension, two teachers mentioned adding worksheets to help with the word study lessons, and two teachers cited needing more opportunities to grade students. Finally, one teacher mentioned adding materials on strategy activities for decoding characters, setting, and plot. Teachers who supplemented generally did so with all of their students, although one teacher only supplemented with ELL students.

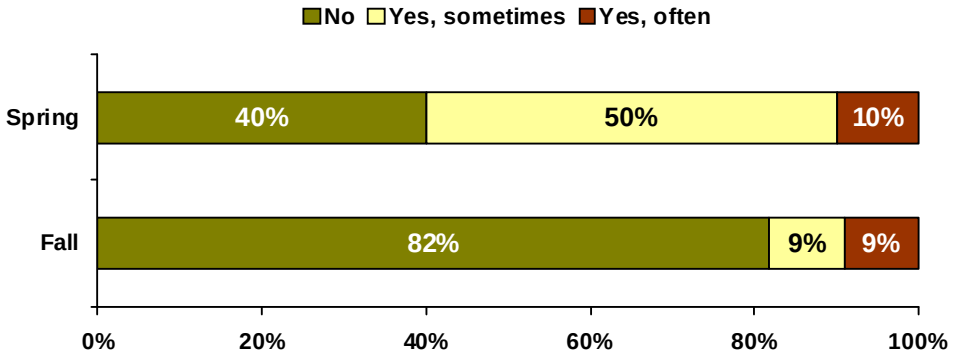


Figure 6. Percentage of teachers who reported using supplemental materials.

Key Question:
What are teachers' perceptions of the quality and utility of the* Literacy by Design *program?

Perceptions of *Literacy by Design*

This portion of the report provides an overview of treatment teachers' perceptions of the *Literacy by Design* program. Evaluators gauged teacher perceptions via online surveys, teacher interviews, and classroom observations.

Ease of Implementation and Utility of Training

The online surveys asked teachers to indicate the relative ease of program implementation. Across both survey administrations, most teachers indicated that the program was not very difficult to implement (see Figure 7). However, teachers generally appeared to find the program easier to

implement over the course of the study. In the spring, the teacher who selected "other" regarding ease of program implementation noted that the program was difficult to implement at first but became much easier over time. The majority of teachers agreed with this statement, reporting on the spring survey that implementation became much easier (70%) or somewhat easier (30%) since the beginning of the year. Discussions with teachers during site visits confirmed the finding that implementation became easier as teachers gained familiarity and comfort with the program.

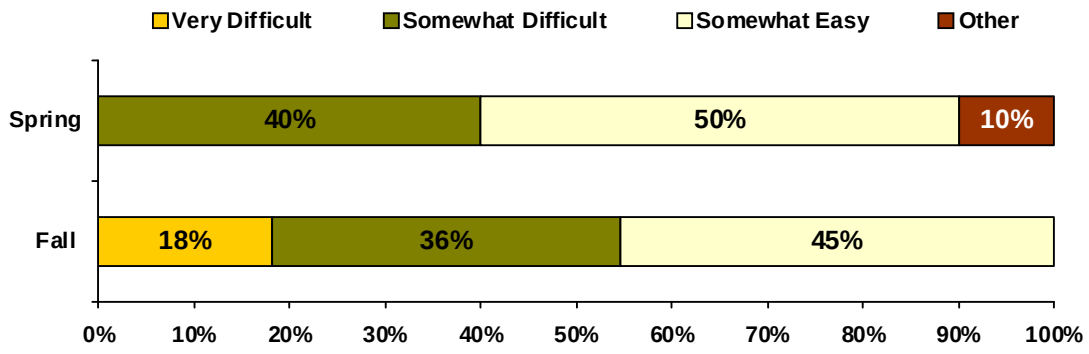


Figure 7. Teacher perceptions regarding ease of program implementation.

Figure 8 indicates that, across both survey administrations, the majority of teachers (95.3%) reported that the *Literacy by Design* training had been very helpful (28.6%) or somewhat helpful (66.7%).

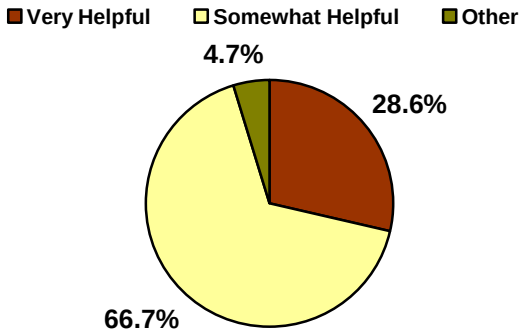


Figure 8. Teachers' perceptions regarding utility of training.

One teacher, who indicated "other" regarding the training, commented that the materials received were helpful. With regard to suggestions for improving training, four teachers mentioned that they would like to see a lesson in use before implementing *Literacy by Design* in their own classrooms. Three other teachers requested more training in the beginning of the year. Along similar lines, two teachers requested more time to use materials and share ideas during the training sessions.

Amount of Materials and Pacing of Instruction

The online surveys asked teachers to indicate their perceptions of the amount of material offered by *Literacy by Design*. On the fall survey, most teachers reported that there was *too much to cover,* but on the spring survey, half of the teachers thought the amount of materials was *just right,* and the remaining teachers thought there was *too much to cover* (40%) or *not enough to cover* (10%). Thus, over time, fewer teachers believed *Literacy by Design* offered too many materials. This finding is not surprising, given that teachers sometimes feel overwhelmed by the amount of materials offered by a new program. As teachers become more comfortable in implementing a program, they often feel less overwhelmed by the amount of materials and instead appreciate the wealth of materials. Indeed, discussions with teachers suggest that they grew to appreciate the amount of materials as they gained experience with the program.

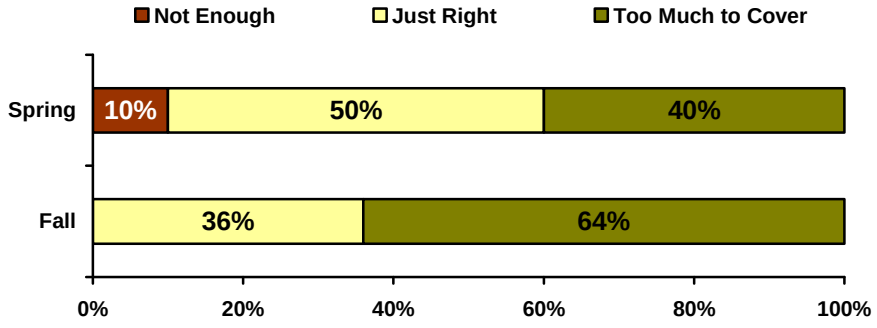


Figure 9. Teachers' perceptions regarding the amount of material offered by *Literacy by Design*.

The online surveys also asked teachers to describe the overall pacing of their instruction. On the fall survey, most teachers (55%; n = 6) indicated that their instruction was fast paced, but by spring, most (60%; n = 6) indicated that their instruction was reasonably paced. Similar to the finding regarding amount of materials, this finding makes sense when one considers that it often takes time to feel comfortable with a new program. Interviews with teachers confirmed that although the initial pacing was somewhat difficult, as teachers gained more experience with the program, the pacing became easier.

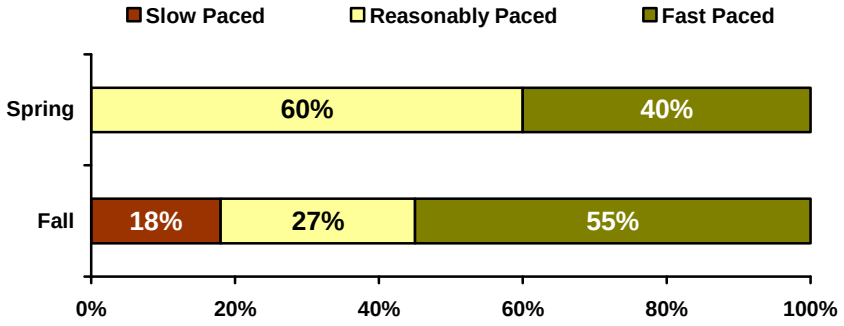


Figure 10. Teachers' perceptions regarding the pacing of the material offered by *Literacy by Design*.

The fall and spring surveys asked teachers to indicate whether the pace of instruction allowed them to meet all of their students' needs. Most teachers (81%; n = 9) agreed that the program allowed them to *meet* or *somewhat meet* all of their students needs. One teacher shared the following comment during an interview:

> *I feel that individual needs are being met more adequately and I know that I am not missing skills or concepts to be taught. I like how it is interwoven with science and social studies.* [Teacher interview, March 2008]

Teachers also had the opportunity to provide feedback regarding the adequacy of the program in meeting the needs of students at different levels. Figure 11 illustrates that most teachers reported that *Literacy by Design* was *adequate* or *very adequate* for allowing teachers to meet the needs of various groups of students. Furthermore, most teachers thought the program was *adequate* or *very adequate*, particularly for advanced and on-level students. Teachers were less likely to view the program as *very adequate* for below and significantly below level students due to the more challenging content present in the program.

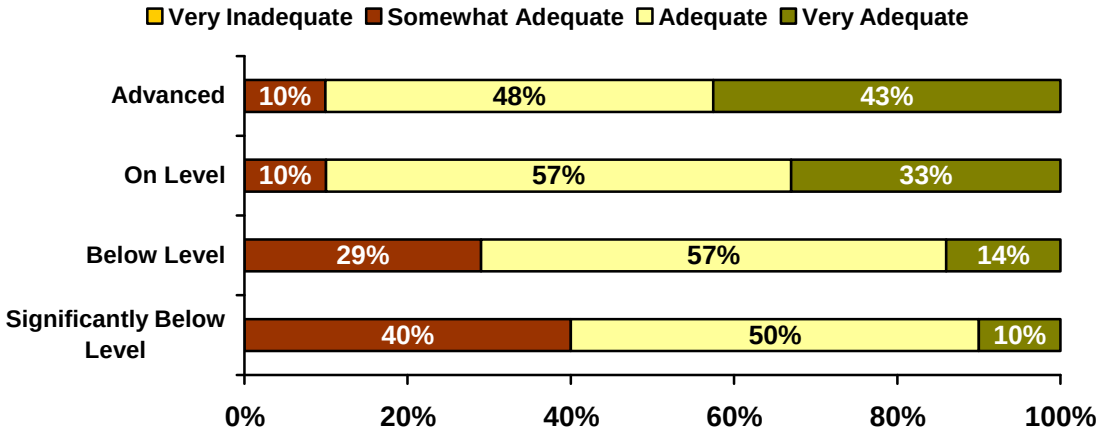


Figure 11. Adequacy of *Literacy by Design* in meeting the needs of students at various levels.

During interviews, teachers mentioned that although the program was most beneficial for on-level and advanced-level students, the small-group instruction was especially beneficial for lower-level readers, because it builds their confidence in reading at their level. Regarding lower-level readers, one teacher shared the following, "the lower students are improving, and they have more confidence in what they are doing. They can figure out what the story is about instead of being frustrated."

Perceptions of Student Learning and Engagement

The spring survey asked teachers to rate their perceptions of the effectiveness of *Literacy by Design* in increasing students learning. The majority of teachers (n = 9) believed the program was somewhat effective to very effective, although one teacher rated the program as *very ineffective* (see Figure 12).

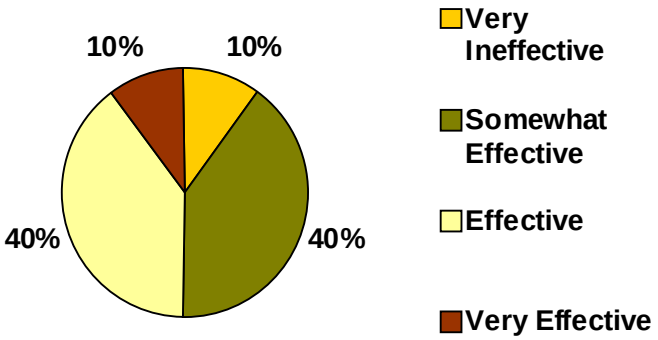


Figure 12. Program effectiveness in increasing student learning across survey administrations.

During interviews, several teachers confirmed that the program had been effective in increasing student learning. One teacher was particularly impressed that even her high-level students made gains and attributed these gains to *Literacy by Design*'s ability to challenge them at their level. Another teacher indicated that the program was really helping the lower-level students and said he had seen more growth with this group of lower students than he had in previous years. Most teachers also indicated that the program had increased their students' interest in reading due to the variety of stories in the program.

Regarding student engagement in the materials, teachers perceived most students to be either averagely engaged in the materials or highly engaged in the materials (see Figure 13).

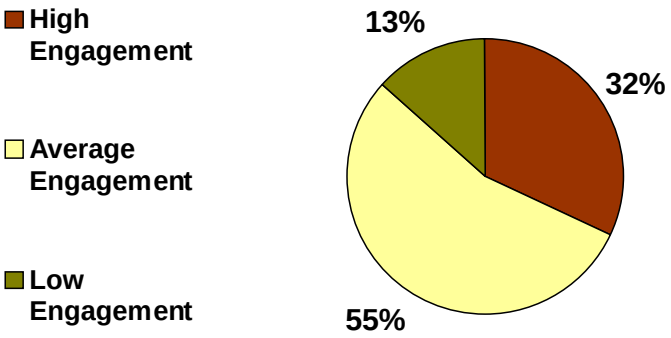


Figure 13. Teachers' perceptions of student engagement across survey administrations.

Findings from classroom observations and interviews with teachers suggested that students enjoyed the program. Teachers commented that students liked the themes and stories and enjoyed participating in small groups. One teacher observed, "There are so many stories that have sparked the interest of especially my boys and my lower level readers."

One teacher shared the following comment about student engagement:

> *There is a lot more interaction and involvement with the children. A lot of the strategies, the think-alongs, and think-alouds are interactive and it makes the children more active participants in the lesson rather than me just presenting.* [Teacher interview, March 2008]

Thus, overall, findings suggest that students were engaged in and enjoyed participating in *Literacy by Design*.

Overall Perceptions of *Literacy by Design*

This portion of the report provides a summary of teachers' perceptions of the *Literacy by Design* program. It also includes feedback from site coordinators.

Perceived Strengths of the Program

Overall, teachers were pleased with the *Literacy by Design* program, and nine out of ten teachers indicated that they wanted to continue using the program next year. One teacher stated, "I love it. I like the laid out lesson plans. I like the themed units, they relate to science and social studies." Another teacher shared the following:

> *I just like how it includes all of the areas of literacy: reading, writing, spelling, vocabulary, grammar. I have had trouble trying to incorporate all of those in an hour and a half. This program makes it flow together and I like that.* [Teacher interview, March 2008]

Teachers praised many aspects of the *Literacy by Design* program, including the following:

- Teachers valued the structure of the program. One teacher commented about how it affected her enjoyment of teaching:

  > *I just enjoyed teaching with the program this year. It really made me a much happier teacher when it came to reading time because I was in control of reading time and the kids knew what we were doing and it was so much more structured and organized for them and they really liked the program a lot.* [Teacher interview, March 2008]

  Two of the sites did not really have structured programs prior to *Literacy by Design*, so they particularly appreciated the structure of the program. Although a few teachers acknowledged that switching to a more structured program was difficult at first (in terms of getting

everything accomplished during class time), they enjoyed the structure and would not want to go back to the unstructured program.

- Teachers appreciated the comprehensiveness of the Teacher's Guides.
- Teachers liked the paradigm of the modeled to independent reading within the lessons and generally thought the program was improving students' reading ability.
- As a group, teachers indicated that the program enabled them to meet the needs of diverse students. They thought the availability of leveled readers was helpful, and they appreciated that their ELLs and special education students could participate in the program fully.
- Several teachers indicated that prior to using *Literacy by Design*, they had not implemented very much small-group instruction in their classroom, and they thought this change was beneficial to students. They also found that students enjoyed working in groups and pairs.

The coordinators also shared a great deal of positive feedback about *Literacy by Design*. Across sites, coordinators indicated that teachers liked the program and wanted to continue using it in the future. Coordinators also generally agreed that *Literacy by Design* enabled teachers to meet students' needs.

Overall, the majority of teachers and coordinators were pleased with the program, citing student and teacher enjoyment in the program. One teacher wrote, "I enjoyed using the program this year. I feel more confident about doing each unit. I like the structure and guidance it provides me. I look forward to using it again next year."

Another teacher added her own perspective, writing the following, "I thought this was a great program. It gave my students a routine they could follow while focusing on the basic skills they need to succeed."

Finally, one teacher added, "I really enjoyed being a part of this study. I feel like I helped more students reach their reading potential with this program than I would have without the program. So thank you very much!"

Perceived Areas for Improvement

Overall, teachers mentioned few areas of improvement for the *Literacy by Design* program. In particular, teachers shared the following feedback:

- One teacher cited the usefulness of the small-group component but was displeased with the whole-group instruction, since it "does not assess comprehension."
- The online survey revealed some teachers were concerned with the difficulty of different aspects of the program, including assessment. One teacher commented

> *I don't like the set-up for the assessment tests. The children are asked to keep changing tasks and those children who cannot change from one task to another find it difficult. Also, the tests are written on grade level and a good portion of the class cannot read on grade level.* [Teacher online survey, October 2007]

Some teachers expressed concerns about not being able to read the themed tests to their students. One teacher explained

> *The themed tests are not good for the low students because they are not able to read the test. I know that the lower kids know more than they are doing. When I helped them they could do it. They've all improved but that has been a problem. Am I testing what they know or if they can read the test?* [Teacher interview, March 2008]

- Initially, a few teachers found the pacing of the program difficult and thought there was not enough time to complete all of the activities, but this seemed to improve as teachers became more comfortable with the program.
- A couple of teachers mentioned that there was too much repetition of the same story within a week and would have preferred to have more time for students to work together.
- Teachers would like to have more professional development for the *Literacy by Design* program, especially regarding pacing of instruction and small-group instruction.
- Several teachers talked about the level of photocopying required with the program and suggested that the program incorporate consumable books instead.
- A small number of teachers indicated that the program was not substitute-teacher friendly because it required synthesizing and other skills that they might not have.
- Some teachers indicated that there were components missing from the whole-class instruction, such as spelling and story sequence. Other teachers commented that the whole-class component does not provide targeted responses or enough information for teachers to know what responses are expected of students.
- Several teachers and coordinators experienced technical difficulties with the *READS* and other online program components. In particular, at the beginning of the year, one site did not feel comfortable using an online assessment and administered the paper-and-pencil version of the *READS*. Another site indicated a desire to access the online fluency component but reported that they were never able to access it.

STUDENT PERFORMANCE RESULTS

To address the study questions regarding the performance of students who participated in the *Literacy by Design* program, evaluators conducted several analyses, including descriptive and inferential analyses, as well as effect size calculations. This section includes a presentation of results for treatment students only, including examination of subgroup learning gains.

LITERACY BY DESIGN IMPACTS ON STUDENT LEARNING

This portion of the report presents the learning gains of students who participated in the *Literacy by Design* program throughout the duration of the study. The report presents results for student performance on the *GMRT-4* and *READS*, broken out by subtest where appropriate.

Key Question:
Do students who participate in the* Literacy by Design *program demonstrate significant learning gains in reading during the study period?

Descriptive Findings

Evaluators conducted descriptive analyses of both the *GMRT-4* and *READS* assessments. This section include findings regarding *GMRT-4* Grade Equivalent Scores, *READS* instructional and independent reading levels, students' proficiency in specific reading skills, and learning gains on the *GMRT-4* and *READS*.

GMRT-4 *Grade Equivalent Scores*

Evaluators examined pretest and posttest *GMRT-4* data to determine whether students participating in the *Literacy by Design* program demonstrated learning gains from the beginning to the end of the study. Figure 18 displays mean pretest and posttest grade equivalent scores on the *GMRT-4*. Findings indicate that, on average, third-grade students who participated in the *Literacy by Design* program improved by 1.7 grade levels on the *GMRT-4* Vocabulary test, 1.4 grade levels on the *GMRT-4* Comprehension test, and 1.5 levels on the *GMRT-4* Total test. Notably, on average, students scored at grade level at pretest, and they scored well above grade level at posttest.

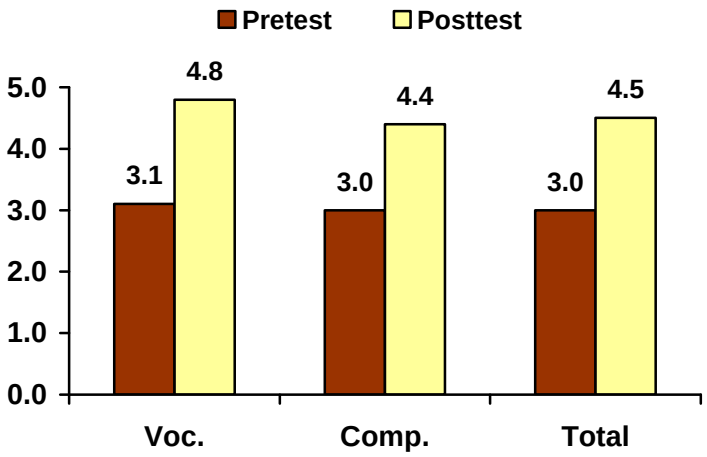


Figure 14. GMRT-4 pre/post grade equivalent scores.

Instructional and Independent Reading Levels

Evaluators also conducted descriptive analyses of the *READS* assessment, which yields information regarding students' instructional and independent reading levels. The instructional reading level is the level at which a student should read for instruction based on *READS* performance. The independent reading level is the level at which a student should be able to read on his or her own without needing instruction. The reading levels for the Grade 3 *READS* test range from 1–4 minus to a 5–1 plus. The first number in the reading level refers to the grade level, and the second number refers to the range within that particular grade level. For example, a score of 3-2 would indicate a reading level of grade three, level 2. Figures 19 and 20 display the percentages of students whose pretest and posttest instructional reading levels correspond to each grade level based on their *READS* scores, and Figures 21 and 22 display the percentage of students whose pretest and posttest independent reading levels correspond to each grade level based on their *READS* scores.

Instructional Reading Level

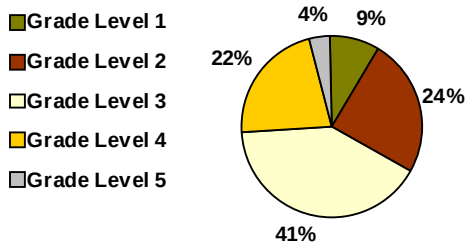


Figure 15. Pretest *READS* Instructional Reading Level.

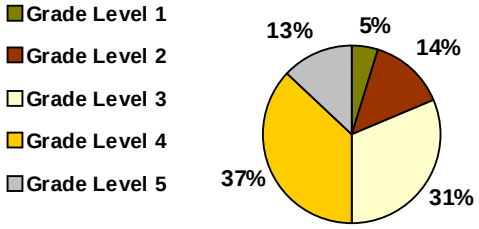


Figure 16. Posttest *READS* Instructional Reading Level.

Independent Reading Level

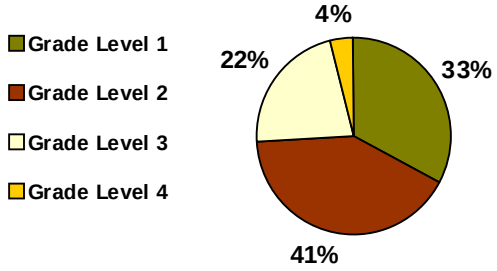


Figure 17. Pretest *READS* Independent Reading Level.

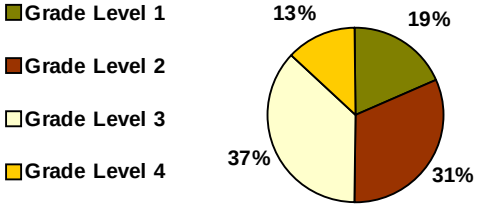


Figure 18. Posttest *READS* Independent Reading Level.

These findings indicate that from pretest to posttest, the percentages of students scoring at the instructional reading levels associated with lower grade levels decreased, and the percentage of students scoring at the instructional reading levels associated with higher grade levels increased. Furthermore, the percentages of students scoring at the independent reading levels associated with lower grade levels decreased, and the percentage of students scoring at the instructional reading levels associated with higher grade levels increased. These findings suggest that students who participated in the *Literacy by Design* program became better able to receive instruction, as well as read independently, at higher levels over the course of the school year.

Proficiency in Specific Reading Skills

The *READS* also yields indicators of whether a student's scores are consistent with a Beginning, Developing, or Proficient classification in Critical Comprehension (CC), Inferential Comprehension (IC), Literary Comprehension (LC), Sounds-Letters Consonants (SLC), Sounds-Letters Vowels (SLV), Vocabulary in Context (VIC), and Word Part Clues (WPC). Figures 23 and 24 display the percentages of students classified as Beginning, Developing, or Proficient in the skill areas assessed at pretest and posttest, respectively. Comparing these figures reveals that, on average, more students moved into the proficient category by posttest for all skills except for Sounds-Letters Consonants.

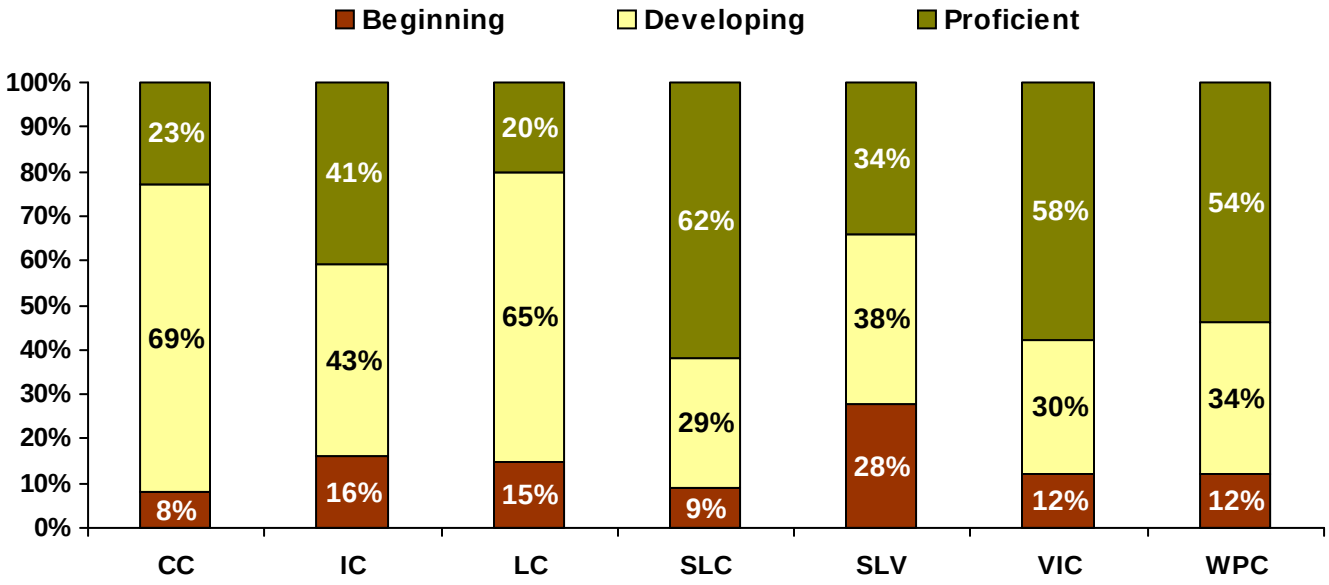


Figure 19. Pretest *READS* Proficiency Levels.

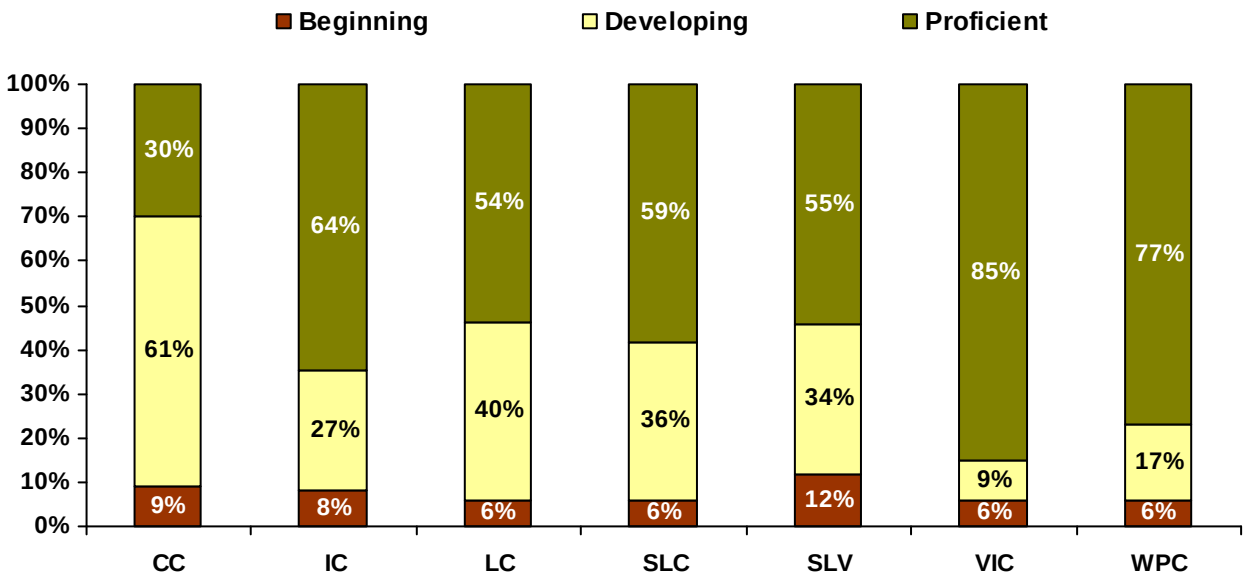


Figure 20. Posttest *READS* Proficiency Levels.

Learning Gains as Evidenced by GMRT-4 *and* READS *Scores*

Evaluators also examined the Extended Scale Scores for the *GMRT-4* Vocabulary, Comprehension, and Total tests, and raw scores for the *READS* Reading Comprehension (RC), Sounds-Letters Consonants (SLC), Sounds-Letters Vowels (SLV), Vocabulary in Context (VIC), and Word Part Clues (WPC) tests and calculated effect sizes to determine the magnitude of changes in scores over the course of the study. Figures 25 and 26 present these data descriptively.

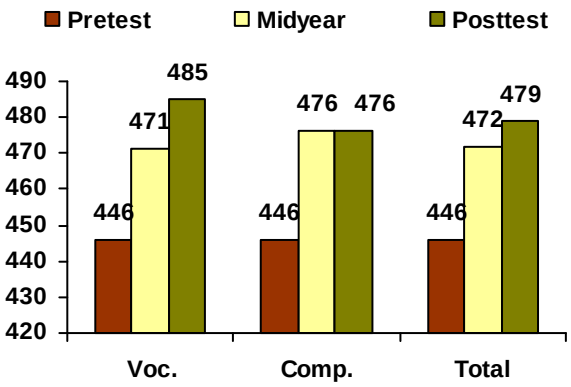


Figure 21. GMRT-4 pre-, mid-, and posttest extended scale scores.

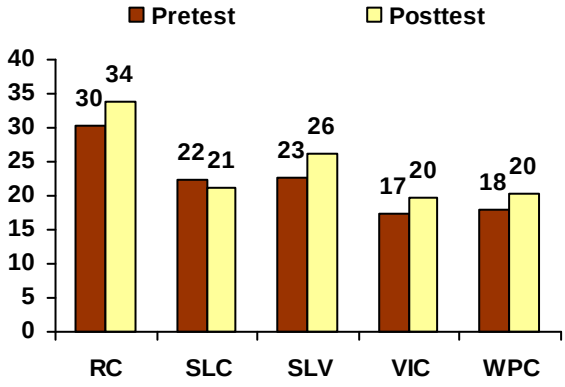


Figure 22. READS pretest and posttest raw scores.

These data indicate that students who participated in the *Literacy by Design* program demonstrated gains from pretest to posttest on the *GMRT-4* Vocabulary, Comprehension, and Total tests. Students participating in the *Literacy by Design* program also demonstrated gains on the *READS* Reading Comprehension, Sounds-Letters Vowels, Vocabulary in Context, and Word Part Clues tests. As a group, they experienced a slight decline on the Sounds-Letters Consonants test. Evaluators conducted multilevel modeling to examine the significance and magnitude of these gains, and these results follow.

Multilevel Analyses

Evaluators conducted multilevel modeling to examine whether or not students who participated in the *Literacy by Design* program demonstrated significant learning gains in reading throughout the study period. Multilevel modeling was appropriate because in this study, the data were nested (i.e., students nested in classrooms). Because these analyses sought to explore the gains of all students who participated in the program (regardless of demographic and background characteristics), the models did not include covariates. Table 4 displays the results, which demonstrate that students participating in the *Literacy by Design* program demonstrated significant gains and large effect sizes on the *GMRT-4*. Specifically, these students demonstrated large gains on the *GMRT* Vocabulary test (d = 1.70), Comprehension test (d = 1.14), and Total test (d = 1.69), translating to 45, 37, and 45 percentile points, respectively.

Table 2.
Mean GMRT Gains for Treatment Students

Outcome Measure	Coefficient	Standard Error	t-value	Approx. df	p-value	Effect Size	Percentile Gain
			GMRT-4				
GMRT-4 Vocabulary	38.88	1.68	23.11	10	0.000*	1.70	45
GMRT-4 Comprehension	30.31	3.73	8.14	10	0.000*	1.14	37
GMRT-4 Total	34.03	2.72	12.50	10	0.000*	1.69	45
			READS				
READS Reading Comprehension	3.48	0.77	4.52	10	0.001*	0.46	17
READS Sounds-Letters Consonants	– 1.01	0.55	–1.83	10	0.100	–0.18	–7
READS Sounds-Letters Vowels	4.03	0.80	5.01	10	0.000*	0.61	23
READS Vocabulary in Context	2.12	0.46	4.59	10	0.001*	0.46	17
READS Word Part Clues	2.58	0.30	8.74	10	0.000*	0.66	24

* Significant at the 0.05 level

On the *READS*, treatment students demonstrated significant gains and medium-to-large effect sizes on four out of the five *READS* subtests. In particular, treatment students demonstrated medium effect sizes ($d = 0.45$) on the *READS* Reading Comprehension and Vocabulary in Context ($d = 0.46$) subtests, translating to gains of 17 percentile points. Treatment students demonstrated large effect sizes ($d = 0.61$) on the *READS* Sounds-Letters Vowels and Word Part Clues ($d = 0.66$) subtests, translating to gains of 23 and 24 percentile points, respectively. Although treatment students demonstrated decreases on the *READS* Sounds-Letters Consonants subtest, these losses were not significant, and the effect sizes were small ($d = -0.18$), translating to decreases of only 7 percentile points.

These results suggest that the *Literacy by Design* program is effective at increasing student learning in reading. Furthermore, many of the gains demonstrated by treatment students were large.

Key Question:
Does participating in the* Literacy by Design *program have differential effects for various subgroups of students and for students in different settings?

Effects of Various Student Characteristics on Student Gains

Evaluators used multilevel modeling to examine the effects of various student characteristics, including gender, ethnicity, free and reduced lunch status, Limited English Proficiency (LEP) status, and special education status on the learning gains of students who participated in the *Literacy by Design* program throughout the school year. The only significant finding among these analyses was that students classified as special education gained less than students not classified as special education on the *READS* Sounds-Letters Consonants subtest did. Because these analyses divide the sample into smaller subsets, caution is warranted when interpreting these results. Therefore, evaluators also calculated effect sizes as a measure of practical significance.

Although the following findings were *not* significant, they are descriptively noteworthy and reflect small-to-medium effect sizes (see Appendix H, Tables H1 and H2):

- Male students gained more than female students did on the *GMRT-4* Vocabulary test ($d = 0.10$). Male students gained less than female students did on the *READS* Reading Comprehension and Sounds-Letters Vowels subtests ($d = -0.15$, $d = -0.15$, respectively), but males gained more than females on the *READS* Sounds-Letters Consonants and Word Part Clues subtests ($d = 0.10$, $d = 0.22$, respectively).
- Caucasian students gained more than ethnic minority students on the *GMRT-4* Vocabulary, Comprehension, and Total tests ($d = 0.22$, $d = 0.40$, $d = 0.33$, respectively). Caucasian students also gained more than ethnic minority students on the *READS* Sounds-Letters Consonants, Sounds-Letters Vowels, and Vocabulary in Context subtests ($d = 0.12$, $d = 0.25$, $d = 0.34$, respectively).
- Students classified as being eligible to receive free or reduced-price lunch gained more than those who were not classified as being eligible to receive free or reduced lunch on the

GMRT Vocabulary and Total tests (d = 0.13 and d = 0.18, respectively). Students classified as being eligible to receive free or reduced-price lunch gained more than those who were not classified as being eligible to receive free or reduced lunch on the *READS* Vocabulary in Context and Word Part Clues subtests (d = 0.23 and d = 0.18, respectively). Students classified as being eligible to receive free or reduced-price lunch gained less than those who were not classified as being eligible to receive free or reduced lunch on the *READS* Sounds-Letters Consonants subtest (d = –0.13).

- Students with limited English proficiency gained more than students who were proficient in English on the *GMRT* Vocabulary, Comprehension, and Total tests (d = 0.25, d = 0.33, and d = 0.34, respectively). Students with limited English proficiency gained more than students who were proficient in English on the *READS* Reading Comprehension, Vocabulary in Context, and Word Part Clues subtests (d = 0.36, d = 0.32, and d = 0.16, respectively).
- Students classified as special education students gained less on the *GMRT-4* Comprehension and Total tests (d = –0.45 and d = –0.46, respectively), and they gained less on the *READS* Reading Comprehension, Sounds-Letters Consonants, Sounds-Letters Vowels, Vocabulary in Context, and Word Part Clues subtests (d = –0.30, d = –0.60, d = –0.13, d = –0.11, and d = –0.50, respectively).

Effectiveness of *Literacy by Design* for Students of Various Ability Levels

More than half of the treatment teachers indicated that they used ability grouping with their students. Therefore, evaluators collected ability grouping data from these teachers to examine whether the *Literacy by Design* program works differentially well for students of varying ability levels. First, evaluators conducted paired samples t-tests to examine whether each ability group (i.e., below-level students, on-level students, and above-level students) demonstrated significant gains on the *GMRT-4* Vocabulary, Comprehension, and Total tests[6].

6 Evaluators conducted similar analyses on the *READS*, and Table H3 in Appendix H displays these results. These findings revealed that within each ability level group, students experienced notable gains corresponding to small, medium, or large effect sizes on all *READS* tests except for the Sounds-Letters Consonants test. It is important to note that the sample sizes in these analyses were somewhat small, so caution is warranted when interpreting the results. Analyses of Variance revealed no significant differences in gains by ability level group on the Reading Comprehension [$F(2,106) = 0.21$, $p = 0.81$], Sounds-Letters Consonants [$F(2,106) = 0.22$, $p = 0.80$], Sounds-Letters Vowels [$F(2,106) = 0.53$, $p = 0.59$], Vocabulary in Context [$F(2,106) = 0.29$, $p = 0.75$], or Word Part Clues [$F(2,106) = 0.96$, $p = 0.39$] tests. These findings indicate that *Literacy by Design* appears to work equally well for improving a number of students' reading skills, regardless of students' initial ability levels.

Table 3.
Paired Samples t-Tests for* GMRT-4 *Gains by Ability Group

Measure	n	Pre/Post Mean Difference	SD	T-value	df	p level	ES (Cohen's *d*)	Percentile Gain
Below-Level Ability Group								
GMRT-4 Vocabulary	46	42.00	24.45	11.65	45	0.000*	1.44	42
GMRT-4 Comprehension	47	27.17	26.09	7.14	46	0.000*	0.98	33
GMRT-4 Total	46	33.17	18.88	11.92	45	0.000*	1.36	41
On-Level Ability Group								
GMRT-4 Vocabulary	39	41.31	19.47	13.25	38	0.000*	1.68	45
GMRT-4 Comprehension	38	33.71	29.62	7.02	37	0.000*	1.14	37
GMRT-4 Total	38	35.58	20.69	10.60	37	0.000*	1.51	43
Above-Level Ability Group								
GMRT-4 Vocabulary	21	33.05	18.41	8.23	20	0.000*	1.09	36
GMRT-4 Comprehension	21	21.48	25.43	3.87	20	0.001*	0.72	26
GMRT-4 Total	21	28.00	21.91	5.86	20	0.000*	1.00	34

*Significant at the 0.05 level

These findings revealed that within each ability level group, students experienced significant gains with large effect sizes on the *GMRT-4* Vocabulary, Comprehension, and Total tests.

Next, evaluators conducted an Analysis of Variance (ANOVA) with the beginning-of-year ability group as the independent variable and gains on the *GMRT-4* Vocabulary, Comprehension, and Total tests as dependent variables. This analysis revealed no significant differences in gains by ability level group on the Vocabulary [$F(2,103) = 1.36$, $p = 0.26$], Comprehension [$F(2,103) = 1.44$, $p = 0.77$], or Total [$F(2,102) = 0.96$, $p = 0.39$] tests. These findings indicate that the *Literacy by Design* program appears to work equally well for increasing students' vocabulary and comprehension skills, regardless of students' initial ability levels.

Evaluators also examined the extent to which ability groups changed over the course of the school year. As Figure 27 illustrates, at the onset of the study, there were more students in the below-level ability group than there were in the on-level or above-level groups. However, by the end of the study, there were more students in the above-level ability group than there were in the below-level or on-level group.

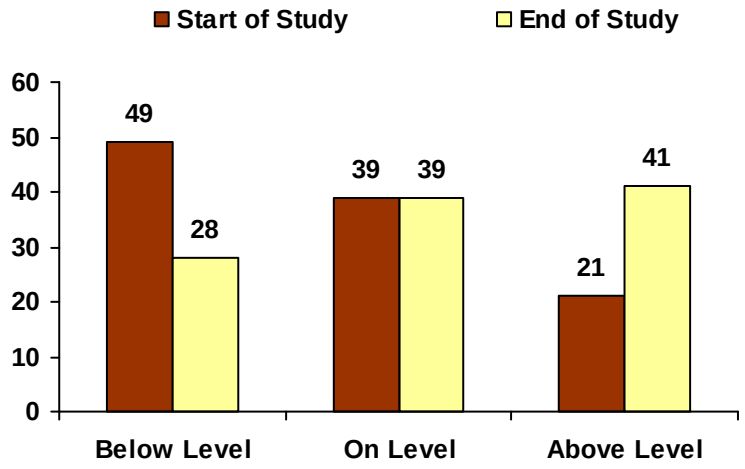


Figure 23. Pre/post changes in numbers of students in each ability group.

Student Performance across Study Sites

To determine whether there were differential learning gains across study sites, evaluators conducted multilevel modeling to examine the main effects of site on *GMRT-4* and *READS* gains. Results revealed significant differences across sites for the *GMRT-4* Vocabulary test, as well as for all of the *READS* subtests. Tables H3 and H4, in Appendix H, display the results of the multilevel analyses. Like other subgroup analyses, the power for these subgroup analyses is less than that for whole-group analyses because subgroup analyses involve looking at only subsets of the data. Therefore, caution is warranted when interpreting these results, and effect sizes might be more meaningful than significance tests in these instances. Notable findings are as follows:

- Site 1 gained significantly more on the *GMRT-4* Vocabulary test than Site 2 did. There were no significant differences by site on the *GMRT-4* Comprehension test, and Site 3 gained significantly more than Site 2 did on the *GMRT* Total test.
- Site 1 gained significantly more than Site 2 did on the *READS* Reading Comprehension Test. Sites 1 and 3 gained significantly more than Site 2 did on the *READS* Sounds-Letters Consonants and Word Part Clues tests. Site 3 gained significantly more than Site 2 did on the *READS* Sounds-Letters Vowels and Vocabulary in Context tests.
- Site 2 gained the least on all assessments, although these differences were not always significant.
- There were no significant differences in gains between Sites 1 and 3.

These findings suggest that the *Literacy by Design* program was equally effective in Sites 1 and 3, but that it may have been less effective in Site 2.

Key Question:
How much improvement do students who participate in the* Literacy by Design *program demonstrate by midyear?

Multilevel Analyses

Evaluators conducted multilevel modeling to examine the degree to which students who participated in the *Literacy by Design* program demonstrated significant learning gains in reading by midyear. Because these analyses sought to explore the gains of all students who participated in the program (regardless of demographic and background characteristics), the models did not include covariates. Table 6 displays the results, which demonstrate that students participating in the *Literacy by Design* program demonstrated significant gains and large effect sizes on the *GMRT-4* by midyear. Specifically, these students demonstrated large gains on the *GMRT* Vocabulary test (d = 1.11), Comprehension test (d = 1.12), and Total test (d = 1.49), translating to 36, 36, and 43 percentile points, respectively. These results suggest that the *Literacy by Design* program is effective at increasing student learning in reading by midyear.

Table 4.
Mean GMRT Gains for Treatment Students

Outcome Measure	Coefficient	Standard Error	t-value	Approx. df	p-value	Effect Size	Percentile Gain
GMRT-4 Vocabulary	25.57	1.96	13.04	10	0.000*	1.11	36
GMRT-4 Comprehension	29.50	2.89	10.21	10	0.000*	1.12	36
GMRT-4 Total	26.79	1.94	13.84	10	0.000*	1.49	43

* Significant at the 0.05 level

CONCLUSION

Literacy by Design is a comprehensive K–5 literacy program that connects whole-class, small-group, and independent instruction through a common comprehension strategy. *Literacy by Design* covers 160 instructional days and is broken out into 16 themes. The program provides direct instruction in phonics, vocabulary, and fluency, and it includes many different opportunities for students to practice skills in these areas. *Literacy by Design* connects instruction to content areas through texts linked to science and social studies themes that correlate to grade-level national standards, and the variety of program materials seek to increase student motivation in reading. *Literacy by Design* offers tools and guidance for ongoing formative and summative assessment, which is critical to informing instruction.

The purpose of this report is to present treatment group findings from a randomized control trial that evaluated teachers' perceptions and implementation of the *Literacy by Design* program and assessed the effectiveness of the materials in helping students attain critical reading skills**.** The final treatment group analytical sample consisted of 11 teachers, three site coordinators, and 199 students across three school districts. For the purposes of this study, treatment teachers implemented *Literacy by Design* instead of their regular core reading programs during the 2007–2008 school year.

The findings indicate that treatment teachers implemented the *Literacy by Design* program appropriately in classrooms. They used it for four to five days per week, incorporated whole-group, small-group, and independent reading instruction, and used all of the core materials. Findings from the online surveys, classroom observations, and interviews revealed that teachers generally liked the *Literacy by Design* program. They really appreciated the wealth of materials, as well as the program's structure, and they found the program enjoyable and engaging for their students. As a group, teachers generally thought the materials were effective at increasing student learning, and all teachers reported that *Literacy by Design* was somewhat adequate to very adequate for allowing teachers to meet the needs of various groups of students. An overwhelming majority of the teachers indicated that they would like to continue using the program next year.

Teachers also provided some constructive feedback about *Literacy by Design*. Some teachers expressed concerns about the assessments, including their difficulty and not being able to read them to their low-level readers. Initially, a few teachers found the pacing of the program difficult and thought there was not enough time to complete all of the activities, but these concerns lessened as teachers became more comfortable with the program. Teachers wanted more professional development for the *Literacy by Design* program, especially regarding pacing of instruction and small-group instruction. Several teachers suggested that the program incorporate consumable books to reduce the level of photocopying required with the program, and a few teachers would have liked to see additional components in the whole-class instruction, such as spelling and story sequence, as well as additional information in the Teacher's Guide so that teachers would know what responses are expected of students. Finally, several teachers and coordinators experienced technical difficulties with the *READS* and other online program components, but even with these difficulties, they still praised the program.

Third-grade students who participated in the *Literacy by Design* program gained more than one grade level in their vocabulary and comprehension reading skills, on average, over course of the study. As a group, treatment students scored at grade level on the *GMRT-4* at the beginning of the study, and they scored well above grade level by the end of the study. These students demonstrated significant large gains on the *GMRT* Vocabulary test ($d = 1.70$), Comprehension test ($d = 1.14$), and Total test ($d = 1.69$), translating to 45, 37, and 45 percentile points, respectively. *Literacy by Design* participants also demonstrated significant gains on the *READS*. In particular, treatment students demonstrated medium effect sizes ($d = 0.45$) on the *READS* Reading Comprehension and Vocabulary in Context ($d = 0.46$) subtests, translating to gains of 17 percentile points. Treatment students demonstrated large effect sizes ($d = 0.61$) on the *READS* Sounds-Letters Vowels and Word Part Clues ($d = 0.66$) subtests, translating to gains of 23 and 24 percentile points, respectively. Treatment group pre/post differences on the *READS* Sounds-Letters Consonants subtest were not significant, and the effect sizes were small ($d = -0.18$), translating to decreases of only 7 percentile points.

In addition to developing specific reading skills, the study's findings suggest that students who participated in the *Literacy by Design* program became better able to receive instruction and read independently at higher levels over the course of the school year. These results suggest that the *Literacy by Design* program is effective at increasing student learning in reading.

Because the classrooms in this study, as well as classrooms nationwide, have students with diverse ability levels, evaluators conducted analyses to examine the effectiveness of *Literacy by Design* with students of varying reading abilities. These findings revealed that within each ability level group (i.e., below-level, on-level, and advanced-level), students demonstrated significant gains with large effect sizes on the *GMRT-4* Vocabulary, Comprehension, and Total tests. Further analyses revealed that there were no significant differences in learning gains by ability group. This important finding suggests that the *Literacy by Design* program works equally well for students, regardless of their ability level. Thus, *Literacy by Design* is appropriate for use in classrooms where students differ in their reading abilities and literacy skills. Additionally, subsequent analyses revealed that many students who participated in *Literacy by Design* moved to higher ability groups over the course of the school year, which confirms earlier findings that the program is effective at increasing students' reading skills.

In conclusion, teachers who participated in the *Literacy by Design* program found it effective and engaging. They appreciated the structure of the program as well as the comprehensiveness of the materials. The findings of this study indicate that during the first year of implementation, *Literacy by Design* was comparably effective with existing programs that skilled teachers had been using for several years. Results reveal that the program is successful in significantly improving children's reading skills in the areas of vocabulary, comprehension, and fluency, as well as helping children become better able to read higher levels of text independently. Furthermore, the program is equally effective for children who read at different levels, making it appropriate for use in classrooms in which students have diverse literacy skills.

REFERENCES

Armbruster, B. B., Lehr, F., & Osborn, J. (2001). *Put reading first: The research building blocks for teaching children to read.* Washington, DC: Partnership for Reading.

Harcourt Achieve (2007). *Scientific research base: Literacy by design.*

Harcourt Achieve (2008). *Introducing literacy by design: Program guide.*

Lou, Y., Abrami, P. C., & Spence, J. (2000). Effects of within-class grouping on student achievement: An exploratory model. *Journal of Educational Research, 94*(2), 101–112.

Mathes, P. G., Torgesen, J. K., Clancy-Menchetti, J., Santi, K., Nicholas, K., Robinson, C., et al. (2003). A comparison of teacher-directed versus peer-assisted instruction to struggling first-grade readers. *The Elementary School Journal 103*(5), 459–480.

National Institute of Child Health and Human Development, NIH, DHHS (2001). *Put reading first: Helping your child learn to read.* Washington, DC: U.S. Government Printing Office.

Neuman, S. B, & Dickinson, D. K. (2001). *Handbook of early literacy research.* New York: Guilford Press.

Pearson, P. D. & Gallagher, M. (1983). The instruction of reading comprehension. *Contemporary Educational Psychology*, *8*.

Raudenbush, S. & Bryk, A. (2002). *Hierarchical Linear Model: Applications and Data Analysis Methods.* Thousand Oaks, CA: Sage Publications.

Raudenbush, S., Spybrook, J., Liu, X. F., & Congdon, R. (2005). *Optimal design for longitudinal and multilevel research.* [computer software and documentation for computer software]. New York, NY: William T. Grant Foundation.

Slavin, R. E., Karweit, N. L., & Madden, N. A. (1989). *Effective programs for students at risk.* Boston: Allyn & Bacon.

The Joint Committee on Standards for Educational Evaluation, Sanders, J. R. (Chair). (1994). *The program evaluation standards* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Tyner, B. (2004). *Beginning reading instruction and the small-group differentiated reading model.* In *Small-Group Reading Instruction* (pp. 1-16). Newark, DE: International Reading Association.

What Works Clearinghouse (2008). *What Works Clearinghouse evidence standards for reviewing studies.* Online. http://ies.ed.gov/ncee/wwc/pdf/study_standards_final.pdf

Appendix A

Study Design and Methodology

In this study, students were nested in classrooms. An appropriate analysis technique for this type of design is hierarchical linear modeling (HLM), also referred to as multilevel modeling, because it accounts for student- and classroom-level variance in student performance outcomes (Raudenbush & Bryk, 2002). Classroom-level variance can be associated with teachers' training and experience, the grade they teach, the reading program they use in their classrooms, and the degree to which they implement the program they use. Student-level variance can be associated with children's gender, ethnicity, socio-economic status, and beginning-of-year reading ability. The idea is that students who share a classroom and teacher also share many of the same classroom-level influences. Therefore, they are more likely to respond similarly to a program than students randomized from different classrooms are (Borman, Slavin, Cheung, Chamberlain, Madden, & Chambers, 2005). Because students' learning experiences within shared classrooms are not independent of each other, data should not be analyzed as such, which makes multilevel modeling an appropriate statistical technique for analyzing nested data.

In addition to using multilevel modeling, evaluators conducted additional analyses. Specifically, evaluators calculated effect sizes to ascertain the magnitude of pretest to posttest changes in students' assessment scores.[7] Furthermore, evaluators conducted descriptive and non-parametric analyses to provide formative feedback to program developers regarding participant characteristics, teacher implementation, and changes in students' reading performance.

While conducting the RCT study, evaluators followed the U.S. Department of Education's What Works Clearinghouse quality standards for research (WWC, 2008) and the Program Evaluation Standards proposed by the Joint Committee on Standards for Educational Evaluation (1994). Thus, the RCT included a treatment group and a control group. Interpretation of the treatment only finding warrants caution.

[7] An effect size is a unit of measurement that expresses the difference in outcome for the average treatment participant from the average control student. It also is used to indicate the strength of the increase or decrease in achievement of students in the treatment group. Effect sizes are expressed in standard deviation units. For example, an effect size of 1.0 would indicate that the mean score in the treatment group was 1.0 standard deviation higher than the mean score in the control group.

Table A1.
Data Analysis Plan Summary for Treatment Group Analyses

Evaluation Question	Analysis
Participant Characteristics	
1. What are the characteristics of students who participated in the study?	Descriptive statistics for grade, gender, ethnicity, free/reduced lunch, migrant status, disabled status, special education status, and limited English proficiency (LEP)
2. What are the characteristics of teachers who participated in the study?	Descriptive statistics for number of years teaching, number of years teaching at current grade, and highest degree
Treatment Student Performance	
3. Do children who participate in the Literacy by Design *program demonstrate significant learning gains in reading during the study period?*	HLM analyses to examine gains in *GMRT* and *READS* scores among treatment students from pretest to posttest. The analyses on *GMRT* scores were conducted on the total Extended Scale Scores for Vocabulary, Comprehension, and Total tests. The analyses on *READS* scores were calculated on raw scores for Reading Comprehension (RC), Sounds-Letters Consonants (SLC), Sounds-Letters Vowels (SLV), Vocabulary in Context (VIC), and Word Part Clues (WPC). The equations for these analyses are as follows: Level-1 equation: $Y_{ij} = \beta_{0j} + r_{ij}$ Level-2 equation: $\beta_{0j} = \gamma_{00} + u_{0j}$ Combining the Level-1 and Level-2 equations yielded estimates of the average *GMRT* and *READS* gains across treatment participants (γ_{00}) and error terms associated with the within and between classroom variabilities (r_{ij}) and (u_{0j}), respectively. HLM5 software for multilevel modeling was used to conduct these analyses (Raudenbush, Byrk, Cheong, Congdon, & du Toit, 2004).
4. Does participating in the Literacy by Design *program have differential effects for various subgroups of students and for students in different settings?*	HLM analyses to examine gains in *GMRT* and *READS* scores among treatment students from pretest to posttest. The analyses on *GMRT* scores were conducted on the total Extended Scale Scores for Vocabulary, Comprehension, and Total tests. The analyses on *READS* scores were calculated on raw scores for Reading Comprehension (RC), Sounds-Letters Consonants (SLC), Sounds-Letters Vowels (SLV), Vocabulary in Context (VIC), and Word Part Clues (WPC). The equations for these analyses are as follows: Level-1 equation: $Y_{ij} = \beta_{0j}$ (student characteristics) $+ r_{ij}$ Level-2 equation: $\beta_{0j} = \gamma_{00}$ (classroom characteristics)$+ u_{0j}$ Combining the Level-1 and Level-2 equations allowed evaluators to examine the impact of student and classroom level characteristics on *GMRT-4* and *READS* gains. In these models, the intercept is specified as random, but the effects of the Level-1 and Level-2 variables are specified as fixed. HLM5 software for multilevel modeling was used to conduct these analyses (Raudenbush, Byrk, Cheong, Congdon, & du Toit, 2004).

Evaluation Question	Analysis
	Level-1 Variables Gender Ethnicity Free and Reduced Lunch Status Special Education Status LEP Level-2 Variables Site
5. To what extent do students experience learning gains by the middle of the school year?	HLM analyses to examine gains in *GMRT* scores among treatment students from pretest to midyear. The analyses on *GMRT* scores were conducted on the total Extended Scale Scores for Vocabulary, Comprehension, and Total tests. The equations for these analyses are as follows: Level-1 equation: $Y_{ij} = \beta_{0j} + r_{ij}$ Level-2 equation: $\beta_{0j} = \gamma_{00} + u_{0j}$ Combining the Level-1 and Level-2 equations yielded estimates of the average *GMRT* gains across treatment participants (γ_{00}) and error terms associated with the within and between classroom variabilities (r_{ij}) and (u_{0j}), respectively. HLM5 software for multilevel modeling was used to conduct these analyses (Raudenbush, Byrk, Cheong, Congdon, & du Toit, 2004).
6. What is the magnitude of the learning gains experienced by treatment students?	Calculate effect sizes for the results from *GMRT* and *READS* analyses using the formula for Cohen's *d* and the formula from Raudenbush et al. (2005) as appropriate. Raudenbush et al. (2005) $\delta = \frac{\gamma_{01}}{\sqrt{\tau + \sigma^2}}$
Teacher Program Implementation	
7. What is the breadth and depth of teachers' use of the Literacy by Design *materials and components?* *8. What are teachers' perceptions and experiences with the materials and components?*	- Descriptive statistics for survey data, as well as content analysis and analytic induction of participant interviews and classroom observations

Appendix B

Data Collection Activities

After completing site recruitment and securing the study's participants, Magnolia Consulting evaluators developed and disseminated evaluation materials to teachers, district coordinators, school coordinators, and materials coordinators. These materials included an overview of the study, a schedule for the study, directions for participating, informed consent forms, and sample data collection instruments. Before teachers began implementing the program, they participated in a study orientation conducted by Magnolia Consulting and a *Literacy by Design* product training facilitated by Houghton Mifflin Harcourt Supplemental. At the onset of the school year, Magnolia Consulting evaluators collaborated with study coordinators to ensure that teachers administered the student assessments during the specified timeframe. Specifically, treatment students completed the *GMRT-4* and the *READS*, and control students completed the *GMRT-4*. The *GMRT-4* and *READS* assessments were scored, and Magnolia Consulting evaluators entered the data into a database for analyses.

Approximately six to eight weeks after program implementation began, Houghton Mifflin Harcourt Supplemental facilitated another training for treatment teachers. During this follow-up training, the training consultant worked with sites to resolve any implementation concerns. Most of these concerns were about not being able to implement every program component during the allotted reading block, and trainers provided guidance, tips, and strategies for teachers. In the fall, Magnolia Consulting evaluators administered an online survey to treatment teachers. This survey asked questions about program implementation and provided teachers with an opportunity to provide feedback about the program. In the winter, evaluators worked with coordinators to oversee the midyear administration of the *GMRT-4*.

In the early spring, evaluators conducted site visits for observations and interviews with all treatment teachers and a sample of control teachers. These site visits enabled evaluators to see how teachers were implementing the program in their classrooms and to learn more about the control group literacy programs. In late spring, treatment teachers completed a second online survey to gauge whether their program implementation and perceptions had changed over the course of the school year. At this time, control teachers also completed an online survey in which they provided information about their literacy instruction. At the end of the school year, treatment students completed the end-of-year *GMRT-4* and *READS* assessments, and control students completed the end-of-year *GMRT-4* assessment.

Appendix C

Assessment Information

Student Assessments

GMRT-4

The *Gates McGinitie Reading Test, Fourth Edition® (GMRT-4)* is a group-administered assessment that gauges decoding skills, vocabulary, reading comprehension, and word knowledge. The Level 3 subtest, which is appropriate for third-grade students, yields Vocabulary, Comprehension, and Total test scores.

The Kuder-Richardson Formula 20 (K-R 20) reliability coefficients for the third grade are 0.96 for both the fall Form S and the spring Form T.

There is also a high correlation between alternate Forms S and T. Reliability coefficients for Levels 2 and 3 indicate strong equivalence between the alternate forms (see Table D1). Total score reliability is 0.95 for second grade and 0.93 for third grade.

Table C1.
***Reliability Information for the* GMRT-4**

Level	Grade	Test	Forms S and T: Correlations *r*	K-R 20: Reliability Coefficients	
				Fall Form S	Spring Form T
3	3	Vocabulary	0.90	0.93	0.94
		Comprehension	0.87	0.93	0.93
		Total	0.93	0.96	0.96

READS

The *Rigby Reading Evaluation and Diagnostic System*™ *(READS)* is a group-administered assessment that gauges reading comprehension, phonemic awareness and phonics, vocabulary, word part clues, skimming and scanning text for information, and fluency. The test yields *Instructional* and *Independent Reading Levels* that indicate the level of text a student can read and comprehend. The *READS* Diagnostic Test for Grade 3 also yields separate scores for Reading Comprehension, Sounds-Letters: Consonants, Sounds-Letters: Vowels, Vocabulary in Context, and Word Part Clues. The *READS* can be administered in an electronic or paper-and-pencil version.

The *READS* includes two forms (A and B), making it appropriate for pretesting and posttesting students. Approximately 93% of the items on the *READS* were drawn from the *Metropolitan Achievements Tests, Sixth Edition (MAT6).*

The Kuder-Richardson Formula 20 (K-R 20) reliability coefficients for each of the subtests for Grade 3 are located in Table D2.

Table C2.
***Reliability Information for the* READS**

Grade	Test	K-R 20: Reliability Coefficients	
		Fall Form A	Spring Form B
3	Sounds-Letters: Consonants	0.84	0.85
	Sounds-Letters: Vowels	0.94	0.90
	Vocabulary in Context	0.89	0.81
	Word Part Clues	0.88	0.83
	Reading Comprehension	0.88	0.88

Teacher Measures

Online Surveys

In the fall, treatment teachers completed an online survey designed to gauge their use of the *Literacy by Design* instructional materials. The survey asked questions about which materials teachers implemented, the extent to which they implemented them, and the nature of their use of materials. The survey also asked questions about teachers' planning and preparation time and gave teachers an opportunity to provide feedback about the program pertaining to the amount of materials offered, the pacing of instruction, the ability to meet students' needs with the program, and student engagement. In the spring, treatment teachers completed another online survey asking many of the same questions as the fall survey but also asking questions about teachers' perceived efficacy of the program, as well as asking if their implementation had changed since the fall survey. Control teachers also completed a spring survey, which asked questions about their reading instruction and gave them an opportunity to provide feedback about the comparison programs. These surveys facilitated evaluators' comparisons of treatment and control group teachers' implementation of reading programs.

Classroom Observation Protocol

Evaluators designed the observation protocol for treatment and control teachers around several main constructs: classroom environment, instructional practices, whole-class reading, small-group reading, independent reading, and student engagement. The protocol was broad enough to capture components of the *Literacy by Design* program as well as those of comparison programs. Each construct had sub-constructs with several checklist items pertaining to each. In observing classroom environment, evaluators accounted for space and room arrangement as well as wall and room displays. In observing instructional practices, evaluators focused on use of assessment and teacher-

student interactions. In observing whole-class reading, small-group reading, and independent reading, evaluators focused on specific strategies and skills associated with each. Finally, in observing student engagement, evaluators focused on students' interest in the lesson as well as the degree to which they stayed on task.

Evaluators used the observation data qualitatively and descriptively to triangulate other data sources, including teacher interviews and online surveys. To ensure inter-rater reliability, evaluators conducted an in-depth debrief of individual ratings (i.e., whether an item was marked as "observed"). This process revealed a high degree of inter-rater reliability.

Interview Protocols

Evaluators developed interview protocols for treatment and control teachers to guide discussions that took place after classroom observations. Interview protocols for treatment teachers focused on the classroom context, material and component use, instructional practices, instructional support, teacher and student impacts, and teacher perceptions of their reading curriculum. Through interviews with a sample of control teachers, evaluators obtained a better sense of their reading instruction. Teacher interviews lasted approximately 20–30 minutes.

Appendix D

Implementation Guidelines

Literacy by Design
Implementation Guidelines

Teachers should implement the following instructional routines during their daily 90-minute reading block:

- Build Reading Skills (whole-group): 20 minutes per day
- Read and Comprehend (whole-group): 20 minutes per day
- Differentiated Reading Instruction (small-group and independent): 45 minutes per day
- Support Reading Independence (whole-group): 5 minutes per day

Build Reading Skills (whole-group, 20 minutes per day):

- Follow the *Comprehensive Teacher's Guide*.
- All components (including those in spelling lists and homework boxes) must be implemented according to the guide using sourcebooks, skills masters, transparencies, and assessment guides.

Read and Comprehend (whole-group, 20 minutes per day):

- Follow the *Comprehensive Teacher's Guide*.
- All core components (including homework, think-aloud, think together, etc.) must be implemented according to the guide using sourcebooks, skills masters, transparencies, and assessment guides.
- Enrichment components (e.g., ELL components, components for struggling readers, enrichment activities, etc.) are optional, depending on students' needs.

Differentiated Reading Instruction (small-group and independent, 45 minutes per day):
Teacher meets with small groups while remaining students work independently.

- **Small-group**:
  - Meet with two groups (20 minutes each) per day.
  - Either continue whole-group lesson in small group using *Comprehension Bridge* or follow small-group lesson plan in *Small Group Reading Teacher's Guide* (depending on students' needs).

- When using the *Comprehension Bridge*, refer to the assessment piece on the back of the bridge to determine when children are ready to move into the lesson plan in the *Small Group Reading Teacher's Guide*.
- When using the *Small Group Reading Teacher's Guide*, you can choose specific components to implement based on students' needs. You do not need to spend two days per book if students are ready to move forward sooner, but students MUST pass the benchmark books before moving forward to the next level. Use oral reading records and fluency assessments (located in Appendix) to monitor student progress.

- **Independent**:
  - During the first 20 minutes, students read independently.
  - During the remaining 20–25 minutes, students practice skills with partners or in small groups.

Support Reading Independence (whole-group, 5 minutes per day):

- Follow the *Comprehensive Teacher's Guide*.

Required Assessment Components:

- **Rigby *READS*:**
  - REQUIRED: Administer at the beginning of the year and again at the end of the year.
- **Theme Progress Tests:**
  - REQUIRED: Administer every other week. Lesson 10 will specify that it should be given.
- **Midyear and End-of-Year Progress Tests:**
  - REQUIRED: Administer the midyear test after week 16 and the end-of-year test after Theme 32.
- **Ongoing Test Practice:**
  - REQUIRED: Lesson 7 will specify that you should use these with students.
- **Benchmark Book Assessment:**
  - REQUIRED: Administer to each student before moving to the next level.
- **Spelling Assessment:**
  - REQUIRED: Administer this in Lesson 5 and Lesson 10 during the reading block time.

Optional Assessment Components:

- **Fluency Readers:**
  - OPTIONAL: can be used during small-group reading
- **Fluency Reader Software:**
  - OPTIONAL: can be used during independent reading time while teachers meet with small groups
- **Comprehension Bridge Rubrics:**
  - OPTIONAL: can be used as needed
- **Independent Reading Self-Assessment:**
  - OPTIONAL: can be used as needed

<u>Supplemental Program Resources</u> (use of these is optional):

- **Fluency software:**
  - OPTIONAL: Students can use this for additional practice during independent reading time.
- **Phonics software:**
  - OPTIONAL: Students can use this for additional practice during independent reading time.
- **Professional development components:**
  - OPTIONAL: Teachers can use these for additional guidance.

***NOTE: The full *Literacy by Design* program also includes a writing component. Use of this component is not required, nor is it being assessed as part of the study. Any use of this component should occur outside of the 90-minute reading block.

Appendix E

Site Characteristics

Table E1.
Site Characteristics by District

	Site 1	Site 2	Site 3
Geographic location and city description*	Middle America; town, remote	Mid-Atlantic/ Northeastern; large city	Midwest; rural, fringe
Total student enrollment	1,618	55,915	2,729
Percent qualifying as low-income	39.9%	47.0%	12.6%
Ethnic breakdown			
Caucasian	93.3%	57.0%	97.6%
African-American	1.5%	16.0%	0.3%
Asian/Pacific Islander	1.5%	8.0%	1.1%
Hispanic	1.9%	20.0%	0.8%
Other	1.8%	0.0%	0.2%
Past performance on statewide assessments	Below average	Average	Above average

Site 1

Site 1 is located in a remote town in Middle America with a population of approximately 4,253. The majority of workers are in manufacturing, retail, and education occupations. According to the 2000 U.S. Census, the estimated median household income was $32,341, and the median age was 38.4 years. The percentage of high school graduates over the age of 25 was 80.9, and the percentage of residents over the age of 25 with at least a bachelor's degree was 12.6.

The school district has six schools, three of which are elementary schools. The elementary schools in this district performed below average on past statewide assessments. Approximately 2.4% of students in this school district were classified as (English Language Learners) ELL for the 2006–2007 school year.

Two schools from Site 1, referred to as Schools A and B, are participating in this study. School A (3–5) is located in a remote town and School B (K–5) is located in a rural town. For the 2007 school year, student enrollment was approximately 259 at School A with a teacher-to-student ratio of 1:12. School B enrolled 94 students with a teacher-to-student ratio of 1:15. Schools A and B qualified as Title I schools for the 2007 school year. Students at School A were predominately Caucasian (92%) followed by African-American (3%). Students at School B were predominately Caucasian (88%), Hispanic (6%), and African-American (3%). At School A, 48.3% were eligible for free/reduced lunch compared to 49.5% at School B. Additionally, 3% of students at School A were classified as Limited English Proficient (LEP) compared to 6% at School B.

Site 2

Site 2 is located in a large Northeastern city with a population of approximately 459,737. The majority of workers are in management, sales, and service occupations. According to the 2000 U.S. Census, the median household income was $55,039, and the median age was 35.9. The percentage of high school graduates over the age of 25 was 82.6, and the percentage of residents over the age of 25 with at least a bachelor's degree was 23.2.

The school district has 60 schools, 42 of which are elementary schools. Teachers at Site 2 reported using a balanced literacy framework. The elementary schools in the district performed average on past statewide assessments. Information on ELL students was not available at the district level.

One school from Site 3, referred to as School C, participated in this study. School C (Grades PK–5) is located in a large city. Student enrollment at School C for the 2006 school year was approximately 940 with a teacher-to-student ratio of 1:14. School C was classified as a Title I school for the 2006 school year. Students at School C were predominately Caucasian (66%), Hispanic (19%), and Asian (9%). Sixty-three percent of students were eligible for free/reduced lunch and 7.2% of students were classified as LEP during the 2006–2007 school year.

Site 3

Site 3 is located in a Midwestern, rural city with an estimated population of approximately 9,435 in 2006. The majority of workers are in education, retail, or manufacturing positions. According to the 2000 U.S. Census, the median household income was $46,938, and the median age was 36.4. The percentage of high school graduates over the age of 25 was 85.4, and the percentage of residents over the age of 25 with at least a bachelor's degree was 19.3.

The school district has four schools, two of which are elementary schools. The elementary schools in the district performed above average on past statewide assessments. No students in this school district were classified as ELL for the 2006–2007 school year.

One school from Site 3, referred to as School D, participated in this study. School D (Grades 3–5) is located in a rural city. Student enrollment at School D for the 2007 school year was approximately 556 with a teacher-to-student ratio of 1:22. School D was not classified as a Title I school. The majority of students at School D were Caucasian (97%) followed by Asian (2%). At School D, 10% were eligible for free/reduced lunch and no students were classified as being LEP for the 2007 school year.

Appendix F

Treatment Student Characteristics

Table F1.
Treatment Student Demographics

Characteristics	Treatment Students (n = 199) Percent	n
Gender		
Male	52.3%	104
Female	47.7%	95
Ethnicity		
African-American	1.5%	3
Hispanic	6.5%	13
Caucasian	86.9%	173
Other	5.0%	10
Free/Reduced Lunch (FRL)		
FRL	40.7%	81
Non-FRL	59.3%	118
Limited English Proficiency (LEP)		
LEP	5.5%	11
Non-LEP	94.5%	188
Special Education		
Special Ed.	7.0%	14
Non-Special Ed.	93.0%	185
Section 504		
Sect. 504	2.0%	4
Non-Sect. 504	98.0%	195
Migrant		
Migrant	0.0%	0
Non-Migrant	100.0%	199

Appendix G

Supporting Tables for Program Implementation

Table G1.
Treatment Teachers' Average use of Whole-class Instruction, Small-group Instruction and Professional Development Materials

	# of teachers using the program					
	1 day per week	2 days per week	3 days per week	4 days per week	5 days per week	Did not use
Whole-class instruction						
Comprehensive Teacher's Guide	-	-	-	1	10	-
Reading transparencies	5	4	-	-	-	2
Sourcebook	-	-	-	2	9	-
Sourcebook audio CDs	-	1	2	-	-	8
Small-group instruction						
Small-group Reading Instruction Teacher's Guide	-	-	1	3	7	-
Leveled Readers	-	-	1	2	8	-
Comprehension Bridge	2	1	2	2	2	2
Assessment Materials						
Assessment Guide	8	2	1	-	-	-
Benchmark Book Evaluation Guide	2	-	-	-	-	9
Benchmark books	2	-	-	-	-	9
READS diagnostic forms	8	-	-	-	-	3
READS Administration Guide	9	-	-	-	-	2
READS Teacher Manual	6	-	-	-	-	5
Professional Development Materials						
Online implementation training	1	-	-	-	-	10
Online book group discussion guides	-	-	-	-	-	11
Lifting Literacy for Academic Achievement	-	-	-	-	-	11
Support materials						
Skills Masters	3	3	3	-	2	-
Comprehension Organizers	3	3	-	2	1	2
Comprehension Strategy Poster	3	3	1	-	2	2
Independent Reading Poster	5	-	-	-	3	3
Technology						
Online Lesson Planner	-	-	-	-	-	11
Developmental Phonics CD-ROM	-	-	-	-	-	11
Fluent Reader Software	-	-	-	-	-	11

Appendix H
Supporting Tables for Student Performance Results

Table H1.
GMRT-4 *Gains by Treatment Student Characteristics*

Outcome Measure	Coefficient	Standard Error	t-value	Approx. df	p-value	Effect Size	Percentile Difference
***GMRT-4* Vocabulary**							
Male (versus female)	2.28	3.27	0.70	194	0.49	0.10	4
Caucasian (versus non-Caucasian)	5.02	4.82	1.04	194	0.30	0.22	8
FRL Status (versus non FRL)	3.06	3.42	0.90	194	0.37	0.13	5
LEP (versus non LEP)	5.73	7.16	0.80	194	0.43	0.25	10
Special Education (versus non SPED)	-1.63	6.58	-0.25	194	0.80	-0.07	2
***GMRT-4* Comprehension**							
Male (versus female)	2.08	3.52	0.59	194	0.56	0.08	3
Caucasian (versus non-Caucasian)	10.60	5.52	1.92	194	0.06	0.40	15
FRL Status (versus non FRL)	1.73	4.08	0.42	194	0.67	0.06	2
LEP (versus non LEP)	8.75	8.22	1.07	194	0.29	0.33	13
Special Education (versus non SPED)	-12.00	7.08	-1.70	194	0.09	-0.45	17
***GMRT-4* Total**							
Male (versus female)	1.68	2.70	0.62	193	0.53	0.08	3
Caucasian (versus non-Caucasian)	6.51	4.23	1.54	193	0.13	0.33	13
FRL Status (versus non FRL)	3.59	3.11	1.15	193	0.25	0.18	7
LEP (versus non LEP)	6.93	6.28	1.11	193	0.27	0.34	13
Special Education (versus non SPED)	-9.34	5.41	-1.73	193	0.09	-0.46	17

* Significant at the 0.05 level

Table H2.
READS *Gains by Treatment Student Characteristics*

Outcome Measure	Coefficient	Standard Error	t-value	Approx. df	p-value	Effect Size	Percentile Difference
***READS* Reading Comprehension**							
Male (versus female)	–1.14	1.13	–1.01	174	0.32	–0.15	6
Caucasian (versus non-Caucasian)	–0.54	1.66	–0.32	174	0.75	–0.07	2
FRL Status (versus non-FRL)	0.30	1.22	0.25	174	0.81	0.04	1
LEP (versus non-LEP)	2.78	2.44	1.14	174	0.26	0.36	14
Special Education (versus non-SPED)	–2.32	2.17	–1.07	174	0.29	–0.30	11
***READS* Sounds-Letters Consonants**							
Male (versus female)	0.47	0.55	0.86	175	0.39	0.10	4
Caucasian (versus non-Caucasian)	0.71	1.23	0.58	175	0.57	0.12	4
FRL Status (versus non-FRL)	–0.75	0.89	–0.83	175	0.41	–0.13	5
LEP (versus non-LEP)	–0.38	1.83	–0.21	175	0.84	–0.07	2
Special Education (versus non-SPED)	–3.38	1.61	–2.10	175	0.04*	–0.60	22
***READS* Sounds-Letters Vowels**							
Male (versus female)	–0.97	0.63	–1.54	175	0.13	–0.15	6
Caucasian (versus non-Caucasian)	1.64	1.41	1.16	175	0.25	0.25	10
FRL Status (versus non-FRL)	–0.42	1.05	–0.40	175	0.69	–0.06	2
LEP (versus non-LEP)	0.60	2.10	0.29	175	0.78	0.09	3
Special Education (versus non-SPED)	–0.89	1.84	–0.49	175	0.63	–0.13	5
***READS* Vocabulary in Context**							
Male (versus female)	–0.07	0.68	–0.11	175	0.92	–0.02	1
Caucasian (versus non-Caucasian)	1.55	0.99	1.57	175	0.12	0.34	13

FRL Status (versus non-FRL)	1.06	0.74	1.44	175	0.15	0.23	9
LEP (versus non-LEP)	1.48	1.48	1.00	175	0.32	0.32	12
Special Education (versus non-SPED)	–0.51	1.31	–0.39	175	0.70	–0.11	4
***READS* Word Part Clues**							
Male (versus female)	0.84	0.58	1.45	175	0.15	0.22	8
Caucasian (versus non-Caucasian)	0.08	0.83	0.09	175	0.93	0.02	1
FRL Status (versus non-FRL)	0.72	0.61	1.17	175	0.24	0.18	7
LEP (versus non-LEP)	0.61	1.23	0.49	175	0.62	0.16	6
Special Education (versus non-SPED)	–1.93	1.11	–1.74	175	0.08	–0.50	–19

* Significant at the 0.05 level

Table H3.
Treatment Group Paired Samples t-Tests for* READS *Gains by Ability Group

Measure	n	Pre/Post Mean Difference	SD	T-value	df	p Level	ES (Cohen's *d*)	Percentile Gain
Below-Level Ability Group								
READS Reading Comprehension	49	3.00	9.48	2.22	48	0.03*	0.33	13
READS Sounds-Letters Consonants	49	–1.35	5.64	–1.67	48	0.10	–0.26	–10
READS Sounds-Letters Vowels	49	2.14	7.67	1.96	48	0.06	0.25	10
READS Vocabulary in Context	49	1.06	6.27	1.86	48	0.24	0.18	7
READS Word Part Clues	49	2.49	4.87	3.58	48	0.001*	0.50	19
On-Level Ability Group								
READS Reading Comprehension	39	2.92	6.29	2.90	38	0.01*	0.56	21
READS Sounds-Letters Consonants	39	–1.28	4.58	–1.75	38	0.09	–0.36	–14
READS Sounds-Letters Vowels	39	3.41	5.71	3.73	38	0.001*	0.53	20
READS Vocabulary in Context	39	1.77	2.19	5.04	38	0.000*	1.26	39
READS Word Part Clues	39	2.67	2.63	6.33	38	0.000*	1.24	39
Above-Level Ability Group								
READS Reading Comprehension	21	4.19	5.06	3.80	20	0.001*	0.92	32
READS Sounds-Letters Consonants	21	–2.19	6.42	–1.57	20	0.13	–0.43	–16
READS Sounds-Letters Vowels	21	3.62	6.57	2.53	20	0.02*	0.41	16

READS Vocabulary in Context	21	1.14	1.74	3.01	20	0.01*	1.00	34
READS Word Part Clues	21	1.33	1.85	3.30	20	0.004*	0.99	33

*Significant at the 0.05 level

Table H4.
Treatment Group* GMRT-4 *Gains by Site

Outcome Measure	Coefficient	Standard Error	t-value	Approx. df	p-value	Effect Size	Percentile Difference
***GMRT-4* Vocabulary**							
Site 1 (versus site 2)	10.07	4.04	2.29	8	0.05*	0.44	17
Site 3 (versus site 2)	2.81	3.83	0.73	8	0.48	0.12	4
***GMRT-4* Comprehension**							
Site 1 (versus site 2)	12.36	9.50	1.30	8	0.23	0.46	17
Site 3 (versus site 2)	11.29	9.21	1.23	8	0.26	0.42	16
***GMRT-4* Total**							
Site 1 (versus site 2)	10.76	6.66	1.62	8	0.14	0.54	20
Site 3 (versus site 2)	7.81	6.40	1.22	8	0.26	0.39	15

* Significant at the 0.05 level

Table H5.
Treatment Group* READS *Gains by Site

Outcome Measure	Coefficient	Standard Error	t-value	Approx. df	p-value	Effect Size	Percentile Difference
***READS* Reading Comprehension**							
Site 1 (versus site 2)	4.24	1.52	2.80	8	0.02*	0.57	21
Site 3 (versus site 2)	2.40	1.45	1.66	8	0.14	0.32	12
***READS* Sounds-Letters Consonants**							
Site 1 (versus site 2)	2.81	1.07	2.62	8	0.03*	0.50	19
Site 3 (versus site 2)	2.62	1.00	2.61	8	0.03*	0.47	18
***READS* Sounds-Letters Vowels**							
Site 1 (versus site 2)	3.22	1.48	2.18	8	0.06	0.51	19
Site 3 (versus site 2)	4.30	0.55	7.78	8	0.000*	0.68	25
***READS* Vocabulary in Context**							
Site 1 (versus site 2)	0.36	0.90	0.40	8	0.70	0.08	3
Site 3 (versus site 2)	2.24	0.85	2.62	8	0.03*	0.50	19
***READS* Word Part Clues**							
Site 1 (versus site 2)	1.82	0.53	3.46	8	0.01*	0.48	18
Site 3 (versus site 2)	1.46	0.30	4.93	8	0.001*	0.38	14

* Significant at the 0.05 level