



THE UNIVERSITY OF
MEMPHIS

The Center for Research in
Educational Policy (CREP)

**The Efficacy of the Leveled Literacy
Intervention System for Students in Grades 3–5:
Data Summary Report for Sandwich Public Schools
2015–2016**





THE UNIVERSITY OF
MEMPHIS™

The Center for Research in
Educational Policy (CREP)

The Efficacy of the Leveled Literacy Intervention System for Students in Grades 3-5: Data Summary Report for Sandwich Public Schools 2015-2016

Dr. Carolyn R. Kaldon
Christine C. Lee
Kathryn Sharpe
Dr. Brenda Gallagher
Sean Holden
Laura Lee
Nikki L. Stevens
Michelle Zhang
Dr. Todd A. Zoblotsky

May 2017

Center for Research in Educational Policy
The University of Memphis
866-670-6147
<http://www.memphis.edu/crep>

Acknowledgements

This study was supported by funding from Heinemann Publishing, publisher of professional resources and a provider of educational services for teachers in kindergarten through college. We appreciate the opportunity to conduct this efficacy study with one of their widely used products -- the Leveled Literacy Intervention system (LLI) -- and thank Sandwich Public Schools for the opportunity to work with their schools, teachers, and students.

For questions about this report or the research study, please contact Dr. Carolyn Kaldon, Principal Investigator, at carolyn.kaldon@memphis.edu or 1-866-670-6147.

Table of Contents

Acknowledgements.....	1
Executive Summary.....	4
Introduction	7
Research Questions	8
Method	8
Setting and Population of Participants	8
Teacher Demographics	9
Instrumentation	12
Student Literacy Achievement.....	13
Intervention Fidelity.....	13
Quality of Core Literacy Instruction	14
School and Home Support for Literacy	15
Stakeholder Feedback.....	16
Instrumentation Summary.....	16
Procedure.....	17
LLI Dosage: Number of Days of LLI Instruction	20
Results.....	20
Student Literacy Achievement.....	21
Benchmark Assessment System (BAS).....	21
Summary	23
Intervention Fidelity.....	24
Leveled Literacy Intervention Observation Tool (LLIOT)	24
Quality of Core Literacy Instruction	30
Literacy Observation Tool (LOT).....	30
School and Home Support for Literacy	33
Leveled Literacy Intervention Teacher Questionnaire – Revised (LLITQ-R).....	33
Classroom Teacher Literacy Instruction Questionnaire (CTLIQ).....	37
Leveled Literacy Intervention Principal Questionnaire (LLIPQ)	41
Home Literacy Support Questionnaire (HLSQ)	41
Stakeholder Feedback.....	43
LLI Teacher Focus Group.....	43
Site Researcher Focus Group	45
Conclusions	46
Limitations of the Study and Recommendations.....	48
Recommendations	49
Design.....	49
Implementation	49
Professional Development (PD) Considerations.....	50
Future Directions	50
References	51
Appendices.....	54
Appendix A: Sandwich Benchmark Level Gain Effect Size Summary	55
Appendix B: Fountas & Pinnell Grade-Level Equivalence Chart	56

List of Tables

Table 1: Demographic Overview of SPS Schools for the 2015-2016 School Year	9
Table 2: Demographic Characteristics of Participating LLI Teachers ($n = 5$)	9
Table 3: Demographic Characteristics of Participating 3rd-5th Grade Classroom Teachers ($n = 9$).....	11
Table 4: Demographic Characteristics of Treatment Group Students ($n = 38$).....	12
Table 5: Demographic Characteristics of Control Group Students ($n = 67$).....	12
Table 6: Summary of Data Sources and Participants by Research Question	16
Table 7: Data Collection Summary	19
Table 8: Control and LLI Gains: Control Students with Baseline Advantage/LLI Students with Outcome Advantage	23
Table 9: LLIOT Response Frequencies ($n = 23$).....	25
Table 10: LLIOT Summary Items ($n = 23$)	26
Table 11: Independent <i>T</i> -Test Results for LLIOT Subscales by Grade Level.....	29
Table 12: LOT Response Frequencies.....	30
Table 13: LLITQ Response Frequencies (Fall 2015 $n = 5$)	33
Table 14: CTLIQ Response Frequencies (Spring 2016 $n = 7$)	37
Table 15: HLSQ Response Frequencies ($n = 48$)	41

Executive Summary

This report summarizes the results of an efficacy study of the Leveled Literacy Intervention system (LLI) conducted by the Center for Research in Educational Policy (CREP) in Sandwich Public Schools during the 2015-2016 school year. Developed by authors Irene C. Fountas and Gay Su Pinnell and published by Heinemann, LLI is a short-term, small-group, supplemental literacy intervention system that uses a series of “leveled” texts (i.e., texts of progressing difficulty) to help students in third through fifth grade achieve grade-level competency in literacy. There were three key purposes of this study: (1) to determine the efficacy of the Leveled Literacy Intervention system (LLI) in increasing literacy achievement for students in grades 3-5 and associated student subgroups; (2) to examine LLI program implementation fidelity in grades 3-5; and (3) to determine perceptions of the LLI system according to relevant stakeholders.

A total of 105 students in grades 3-5 participated in this mixed-methods quasi-experimental study (QED) that included both quantitative and qualitative data. The students were matched demographically and assigned to treatment and control groups. During the study, the treatment group students participated in LLI (approximately 24 weeks), while the control group students could not receive LLI until after the study was over. The control students could receive other literacy interventions, however. Treatment and control group students’ pre- and posttest performance was compared on two measures of student literacy achievement: the Fountas & Pinnell Benchmark Assessment System, and their state’s reading assessment. Further, an assessment of LLI implementation fidelity included independent observations of LLI groups and teacher-provided data taken from the LLI Online Data Management System. The quality of the core literacy instruction was also examined using classroom literacy observations, and feedback regarding LLI and the participating schools’ core literacy programs was obtained from LLI teachers, classroom teachers, principals, parents/guardians, and independent site researchers who collected data for the study. Results from the current study are summarized by research question below.

1. What progress in literacy achievement, if any, do students who receive LLI in grades 3-5 make compared to students who receive core literacy instruction alone?

Results revealed that LLI had a positive impact on some 3rd-5th grade students’ literacy achievement. In particular, one of the three types of analyses showed important results: When the control group had an advantage at baseline and the treatment still had an advantage in outcomes. While none of these positive effects were statistically significant, some were considered substantively important based on guidelines from the What Works Clearinghouse (WWC) (i.e., an effect size of +/- 0.25; What Works Clearinghouse, 2014). Detailed findings for each subgroup are presented in the body of this report; however, some of the strongest findings are outlined below.

With regard to **benchmark levels**:

- When starting at a disadvantage, 3rd grade LLI students overall showed substantively higher gains compared to control students.
- When starting at a disadvantage, 4th grade LLI students overall, as well as high achieving students, showed substantively higher gains compared to control students.
- In addition, when starting at a disadvantage, 3rd grade high achieving LLI students and 5th grade LLI students overall showed higher gains compared to control, but these did not qualify as substantively important.

2. At what level of fidelity to the program model is LLI implemented by teachers participating in the study?

Overall, the observation results from the current study suggest that LLI was implemented with inconsistent fidelity to design. When observed, many lesson components received acceptable to high fidelity ratings, with few indications of needing improvement. However, a number of components went unobserved in over half of the observations. In addition, a few areas were rated as needing improvement, namely introducing vocabulary words (30.4%), having well organized lessons (34.8%), and appropriately pacing lesson components (56.5%). Further, the majority of site researchers concluded that, overall, the lessons needed improvement, as they were not delivered as designed (60.9%). However, site researchers had a more neutral opinion in their open-ended comments, of which very few indicated specific areas needing improvement. Additionally, the observation results revealed that LLI implementation was mostly consistent over the school year at both time points when the observations were conducted. Changes in implementation over the year were only noted for 5th grade, where there was substantively important improvement in quality of LLI instruction for 5th grade but slight decreases in the areas of literacy instructional strategies and learning environment. In general, it should also be noted that, for 3rd grade, all subscales were rated between needing improvement and acceptable at both time points.

The observation results were complemented with self-report feedback from the participating LLI teachers, which showed a slightly different picture. In particular, a large majority reported implementing LLI as designed (e.g., meeting daily for 45 minutes, following the LLI Lesson Guide), understanding the LLI goals and procedures, and having sufficient training to implement LLI effectively. However, LLI teachers did note that they were often asked to do other tasks that conflicted with LLI lesson time, feeling that they did not have sufficient planning time to implement the program and that administration did not always protect the time needed for uninterrupted teaching. This may have impacted teachers' implementation and students' progress as seen in the overall achievement results.

Finally, the LLI attendance records that were available (97% of treatment group) from the current study revealed that there were, on average, 67 LLI instructional days, compared to the recommended number of LLI instructional days (i.e., approximately 90-120 days/18-24 weeks). Of these students with attendance data, when looking individually at each student, the data revealed that none of the treatment students received the recommended dosage. Student absences were due to several student-level factors (e.g., individual absences or unavailability during LLI group time) as well as school or district limitations (e.g., holidays, assessment windows during which LLI teachers and/or students were pulled during LLI group time, delays in starting LLI due to scheduling conflicts or difficulty accessing student data). Therefore, the findings at each grade level which are not meeting statistical significance or substantively important progress may have been impacted due to a large number of treatment students not receiving a full dosage of LLI. Schools should note the importance of consistently providing LLI throughout the year so the students can make the most progress by receiving, at a minimum, the recommended amount of LLI lessons.

3. What are stakeholders' perceptions of the LLI system for grades 3-5 and the core literacy program?

Overall, LLI teachers, classroom teachers, principals, parents/guardians, and site researchers shared positive perceptions of the LLI system and its impact on struggling students' literacy success. Stakeholders felt that LLI has benefits for students' literacy achievement and skills as well as their

engagement, interest, and confidence related to reading and writing. Stakeholders also reported positive perceptions of such aspects of the LLI system as its design, instructional components, and materials (particularly the lesson books). However, although stakeholders generally perceived LLI as helpful, there was common feedback regarding the need for improving the logistics of implementing LLI. In general, stakeholders discussed the need for more time to complete lessons and for that time to be protected for LLI teaching, as well as needs for better scheduling and more staff in order to appropriately serve students.

Regarding the core literacy instruction, stakeholders' perceptions were generally positive, although some areas of concern were identified. Stakeholders perceived that their schools are generally supportive of literacy and provide a high-quality learning environment conducive to learning. Further, stakeholders shared positive perceptions of the core literacy program's impact on student engagement, as well as such aspects of the program as classroom materials. The presence of guided reading, independent literacy work, and small, as well as whole group instruction were also noted as being positive aspects of the core literacy programs. However, stakeholders agreed that the core literacy instruction also has areas of improvement. Areas for improvement suggested by stakeholders included more time for literacy, better scheduling, clearer communication of details about the core literacy program, and an increased emphasis on technical reading and writing skills, including grammar.

The current study encountered several limitations that may limit the generalizability of the findings and that prevented researchers from obtaining adequate power to draw definitive conclusions. These limitations included primarily that the sample was not randomized, which, while not ideal for research, was a real-world constraint for obtaining districts that would participate in the study. Other limitations include that none of the treatment group students received the recommended amount of LLI instructional time and the acknowledgement that control group students were allowed to receive other supplemental literacy services besides LLI while they were participating in the study. However, despite these limitations, the current study found educationally meaningful, positive effects of LLI on students' literacy achievement when implemented with sufficient fidelity to the LLI model. Further, stakeholders in these districts – including teachers, administrators, and parents/guardians – were supportive of LLI and perceived positive benefits of the LLI system for their students. Altogether, the results from this research allow us to conclude that LLI has a positive impact on some 3rd-5th grade students' literacy achievement, particularly in 3rd and 4th grades. These results also suggest that continued implementation of LLI would be beneficial in each of these three participating districts and offer an opportunity for research-based recommendations that may enhance the system, future research, and ultimately student achievement. A list of these recommendations – including items related to LLI design, implementation, and professional development, as well as future directions for LLI research – may be found in the main body of this report.

Introduction

This report summarizes the results of a multi-site efficacy study of the Leveled Literacy Intervention system (LLI) conducted by the Center for Research in Educational Policy (CREP) in Sandwich Public Schools (SPS) during the 2015-2016 school year. The school within this district has widely adopted the targeted, small-group implementation model of LLI with support from Heinemann consultants providing LLI professional development, along with continuing support and development provided by trained staff in each district. This report focuses on the implementation and impact of the LLI System for grades 3-5 in an elementary school that voluntarily adopted the LLI system.

CREP is a State of Tennessee Center of Excellence, located at the University of Memphis, whose mission is to implement a research agenda associated with educational policies and practices in preK-12 public schools and to provide a knowledge base for use by educational practitioners and policymakers. Since 1989, the Center has served as a mechanism for mobilizing community and university resources to address educational problems and to meet the University's commitment to primary and secondary schools. Functioning as a part of the College of Education, CREP seeks to accomplish its mission through a series of investigations conducted by Center personnel, college and university faculty, and graduate students.

This study was designed to extend the findings of prior LLI research conducted by CREP regarding the efficacy of LLI in grades K-2 in rural, suburban, and urban settings. While the efficacy of LLI was established in the prior study for the students in these settings, additional research was needed to establish LLI's effectiveness in grades 3-5 across multiple settings. Reading is the tool for learning knowledge in the upper-grade classroom, which plays a central role for academic success across different subjects (Lubliner, 2004; Salinger, 2003). In the long run, without effective reading intervention, struggling upper-graders are likely to experience frustration and failure when they move into middle school and beyond (Lubliner, 2004).

Schools in both rural and urban areas often face challenges such as limited resources, issues related to student mobility and teacher retention, and students often come from high-risk neighborhoods into high-risk schools. Further, these at-risk school districts have a great need for research-based programs that clearly demonstrate a positive impact on student achievement. The goal of this study was to examine the extent to which participation in LLI influenced student literacy achievement and teachers' instructional practices in LLI. Additionally, this study was designed to determine the strengths and weaknesses of LLI according to relevant stakeholders.

Research Questions

There were three key purposes of this study: (1) to determine the efficacy of the Leveled Literacy Intervention system (LLI) in increasing literacy achievement for students in grades 3-5 and associated student subgroups; (2) to examine LLI program implementation fidelity in grades 3-5; and (3) to determine perceptions of the LLI system according to relevant stakeholders. The study used a mixed-methods design to address the following confirmatory and exploratory research questions:

1. What progress in literacy achievement, if any, do students who receive LLI in grades 3-5 make compared to students who receive core literacy instruction alone?
 - a. Does the effectiveness of LLI vary by the following subgroups: English Language Learners, students with a special education designation, and ethnic minorities (i.e., African-American and Hispanic students)?
2. At what level of fidelity to the program model is LLI implemented by teachers participating in the study?
3. What are stakeholders' perceptions of the LLI system for grades 3-5 and the core literacy program?

Method

The present study of the LLI system employed a mixed-methods quasi-experimental design that included both quantitative and qualitative data. A matched-pair design was utilized to assist in the equivalency between treatment and control groups, and pre-post comparisons of student achievement in literacy were conducted. In addition, an assessment of fidelity of implementation – including LLI and classroom observations as well as feedback from teachers, parents/guardians, school-level administrators, and independent site researchers – yielded both observational and self-reported data.

Multiple instruments were utilized in the study, including two measures of reading achievement for evaluating students' progress in literacy; two observational tools for assessing LLI and classroom teachers' instructional practices; and four surveys – along with two focus groups – to obtain feedback on LLI and the core literacy program from LLI and classroom teachers, parents/guardians, principals, and site researchers. Details of each instrument will be discussed later in this section.

Setting and Population of Participants

Two elementary schools serve Sandwich Public Schools (SPS) in Sandwich, Massachusetts, and one was eligible to participate in the study. SPS is a suburban district that served 2,773 students and employed 216 teachers during the 2015-2016 school year. There are four schools in the district, including two elementary schools and two middle/high schools. The majority of students are White (94.8%) with one eighth of the students enrolled in free/reduced lunch (13.1%). Additionally, Special education students make up approximately one sixth of the student population (15.7%) while English Language Learners less than one percent (0.4%). Table 1 summarizes the overall demographic characteristics of the district.

Table 1: Demographic Overview of SPS Schools for the 2015-2016 School Year

Grade Levels	District-Wide Population			District-Wide Student Demographics							
	Students	Teachers	Student/Teacher Ratio	% Asian	% Black	% Hispanic	% White	% Other	% Free/Reduced Lunch	% Special Education	% English Language Learners
ECE-12	2,773	216	12.8	2.2	1.1	1.0	94.8	0.3	13.1	15.7	0.4

Source: Massachusetts Department of Elementary and Secondary Education School and District Profiles (<http://profiles.doe.mass.edu/profiles/student.aspx?orgcode=02610000&orgtypecode=5&>).

Teacher Demographics

When asked to respond to a survey for LLI teachers, data was obtained from 5 participating LLI teachers. The majority of LLI teachers in the study taught 4th grade, followed by 3rd grade, and then 5th grade. Almost half of the teachers had been at their current school for one to five years and over half had been teaching for 15 or more years. In addition, over half of participating LLI teachers had obtained a Master’s degree or Master’s plus 30, while the majority of LLI teachers had acquired a regular/professional teaching certificate and all were fully trained in LLI. All of the teachers were female, and all were White. Overall, these teachers had a solid background of teaching experience at their current school and teaching in general, and half of them had pursued Master’s degrees and taken additional hours beyond their Master’s. Taken together, they appear to have been well positioned to implement the LLI curriculum. Table 2 summarizes the demographic characteristics of the LLI teachers in the study, as reported on the LLI teacher survey.

Table 2: Demographic Characteristics of Participating LLI Teachers (n = 5)

Item	Percent Responded
Grade level(s) taught	
3 rd grade	40.0
4 th grade	60.0
5 th grade	20.0
Years of teaching experience at current school	
Less than 1 year	40.0
1-5 years	40.0
6-10 years	0.0
11-15 years	0.0
More than 15 years	20.0
Years of teaching experience at any school	
Less than 1 year	0.0
1-5 years	0.0
6-10 years	20.0
11-15 years	20.0
More than 15 years	60.0
Highest level of education completed	
Bachelor’s degree	40.0
Master’s degree	40.0
Master’s plus 30 hours	20.0
Education Specialist degree	0.0
Doctoral degree	0.0

Table 2: Continued

Item	Percent Responded
Ethnicity	
Asian or Pacific Islander	0.0
American Indian or Alaskan Native	0.0
African-American/Black	0.0
Hispanic	0.0
White, not of Hispanic origin	100.0
Multi-racial/Other	0.0
Gender	
Male	0.0
Female	100.0
Age group	
29 years or less	0.0
30-39 years	20.0
40-49 years	0.0
50-59 years	40.0
60 years or older	40.0
Level of LLI training	
Completed training	100.0
Partially trained	0.0
None	0.0
Teacher certification level	
Paraprofessional	20.0
Alternative certificate	0.0
Initial/apprentice certificate	0.0
Regular/professional certificate	80.0

Note. Item percentages may not total 100% due to missing input or multiple responses from some participants.

Nine SPS teachers (2 in the Fall, 7 in the Spring) participated in the classroom teacher survey. According to data obtained from this survey, the classroom teachers in the current study primarily taught 5th grade, followed by 4th grade, and then 3rd grade. Slightly over half of the classroom teachers had been at their current school for more than 15 years, and the majority had been teaching in general for over 15 years. Further, the majority held a Master’s degree or Master’s plus 30 hours, and all held a regular/professional teaching certificate. All of the participating classroom teachers were female, and all were White. Overall, the participating classroom teachers were generally well-qualified and had a substantial amount of teaching experience. Table 3 summarizes the demographic characteristics of the 3rd-5th classroom teachers in the study, as reported on the classroom teacher survey.

Table 3: Demographic Characteristics of Participating 3rd-5th Grade Classroom Teachers (n =9)

Item	Percent Responded
Grade level(s) taught	
3 rd grade	14.3
4 th grade	39.3
5 th grade	46.5
Years of teaching experience at current school	
Less than 1 year	21.5
1-5 years	0.0
6-10 years	21.5
11-15 years	0.0
More than 15 years	57.2
Years of teaching experience at any school	
Less than 1 year	0.0
1-5 years	0.0
6-10 years	7.2
11-15 years	7.2
More than 15 years	85.7
Highest level of education completed	
Bachelor's degree	14.3
Master's degree	14.3
Master's plus 30 hours	71.5
Education Specialist degree	0.0
Doctoral degree	0.0
Ethnicity	
Asian or Pacific Islander	0.0
American Indian or Alaskan Native	0.0
African-American/Black	0.0
Hispanic	0.0
White, not of Hispanic origin	100.0
Multi-racial/Other	0.0
Gender	
Male	0.0
Female	100.0
Age group	
29 years or less	0.0
30-39 years	7.2
40-49 years	0.0
50-59 years	53.6
60 years or older	39.3
Teacher certification level	
Paraprofessional	0.0
Alternative certificate	0.0
Initial/apprentice certificate	0.0
Regular/professional certificate	100.0

Note. Item percentages may not total 100% due to missing input or multiple responses from some participants.

Student Demographics

Treatment Group. Within the participating school in SPS, there were a total of 38 students who comprised the treatment group for the study. 100% of treatment students in the sample were White. Over half of the students were male (63.2%). In addition, 10.2% had a special education designation. SPS differed from the overall treatment group sample in many areas, with all of their treatment students being White. Table 4 summarizes the demographic characteristics.

Table 4: Demographic Characteristics of Treatment Group Students (n =38)

District	Grade Levels	Students	% Asian	% Black	% Hispanic	% White	% Multi/ Other	% Male	% Female	% English Language Learners	% Special Education	% Economically Disadvantaged
SPS	3-5	38	0.0	0.0	0.0	100.0	0.0	63.2	36.8	0.0	10.5	21.1

Source: Student records provided by the SPS School Records Department.

Control Group. Across the participating schools in SPS there were a total of 67 students within the control group. Within the participating school, 97% were White, 1.5% were Asian, and 1.5% were of another or mixed ethnicity. Slightly over half of the students were female (56.7%). In addition, 7.5% had a special education designation. SPS differed from the overall control group sample in many areas, with the majority of their control students being White. Table 5 summarizes the demographic characteristics overall and each district separately.

Table 5: Demographic Characteristics of Control Group Students (n =67)

District	Grade Levels	Students	% Asian	% Black	% Hispanic	% White	% Multi/ Other	% Male	% Female	% English Language Learners	% Special Education	% Economically Disadvantaged
SPS	3-5	67	1.5	0.0	0.0	97.0	1.5	43.3	56.7	0.0	7.5	13.4

Source: Student records provided by the SPS School Records Department.

Instrumentation

Both quantitative and qualitative data were collected in this efficacy study. CREP researchers used three measures of reading achievement for evaluating students' progress in literacy: the Fountas & Pinnell Benchmark Assessment System (BAS) was administered by local site researchers trained by CREP, while the state assessments in each district are routinely administered by the district each spring. Two observational tools developed by CREP – the Leveled Literacy Intervention Observation Tool (LLIOT) and the Literacy Observation Tool (LOT) – were used to evaluate LLI and classroom teachers' literacy practices and instructional strategies in the classroom. CREP also developed two teacher surveys, the Leveled Literacy Intervention Teacher Questionnaire – Revised (LLITQ-R) and the Classroom Teacher Literacy Instruction Questionnaire (CTLIQ), as well as a principal survey (the Leveled Literacy Intervention Principal Questionnaire, or LLIPQ) and a parent/guardian survey (the Home Literacy Support Questionnaire, or HLSQ), to ascertain these stakeholders' feedback on LLI and core literacy classroom instruction. Finally, structured focus groups were conducted with LLI teachers and site researchers to gather additional qualitative feedback regarding LLI. Details of each instrument are discussed below.

Student Literacy Achievement

Fountas & Pinnell Benchmark Assessment System (BAS)

The Fountas & Pinnell Benchmark Assessment System 1, 2nd Edition (2010) was used to measure the following literacy skills: phonemic awareness, letter-sound relationships (decoding), vocabulary, comprehension, fluency, and writing. Both treatment and control group students in the study were tested by independent site researchers at the beginning and end of LLI. These data were used to measure individual student gains as well as the composition of the groups in respect to homogeneity of student needs.

The Fountas & Pinnell Benchmark Assessment System (BAS) is an individually administered assessment tool designed by the developers of LLI to reliably place students on the Fountas & Pinnell Text Level Gradient™ (Fountas & Pinnell, 2007), an A-Z gradient of text difficulty. LLI is comprised of three systems for upper elementary grades: Levels L-Q are in the Red System; Levels O-T are in the Gold System; and Levels R-W are in the Purple System. The Red System is generally used in 3rd grade, the Gold System in 4th grade, and the Purple System in 5th grade. The goal of the LLI system is to bring children up to their current grade level in reading, starting from the earliest Level A (usually mid-kindergarten) to Level W (early 4th grade). System 1 of the BAS, which is designed for students reading Levels A-N, and System 2 of the BAS, which is designed for students reading Levels L-Z, use both fiction and nonfiction texts to determine an independent and an instructional reading level for the student. The BAS demonstrates high test-retest reliability (0.97 overall), and convergent validity was established between the reading accuracy rates of BAS System 1 books and those of Reading Recovery assessment texts (0.94 for fiction, 0.93 for nonfiction; Heinemann, 2007). For BAS System 2 books, convergent validity was moderately established with the Slossen measure of word reading (0.69 for fiction, 0.62 for nonfiction) and the Degrees of Reading Power text passage reading assessment (0.44 for fiction, 0.42 for nonfiction; Heinemann, 2007).

State Assessments in Literacy

As a second measure of student literacy achievement, we also requested state assessment scores and proficiency levels in literacy from each of the three participating districts. In SPS, the MCAS is administered each Spring in grades 3-5, and we requested pretest (Spring 2015) and posttest (Spring 2016) scores in literacy for each participating 4th and 5th grade study student. Additional technical information on the MCAS assessment can be found on <http://www.doe.mass.edu/mcas/tech/>.

Intervention Fidelity

Leveled Literacy Intervention Observation Tool (LLIOT): Grades 3-5

The Leveled Literacy Intervention Observation Tool (LLIOT), developed by CREP researchers for a previous study of LLI, involves a targeted, 30-minute observation of an LLI group completing a randomly selected LLI lesson. The LLIOT is used to rate LLI teachers' fidelity to the LLI model as well as the quality of their literacy instructional strategies and the learning environment of the lesson. Ratings are provided using a 4-point scale that ranges from 0 (Not Observed) to 3 (Excellent). Containing 32 items, the LLIOT is comprised of three subscales: Quality of LLI Implementation (19 items), which is designed to measure LLI teachers' implementation of the 10 main LLI lesson components; Literacy Instructional

Strategies (6 items), which is designed to assess LLI teachers' use of general teaching strategies that should be present in a successful literacy intervention; and Learning Environment (7 items), which is designed to assess the quality of lesson factors such as organization, pacing, and the availability of materials.

Site researchers trained by CREP conducted observations of two intervention sessions with each participating LLI group, one near the beginning of LLI and one near the end, using the LLIOT. This observation data contributed to the evaluation of fidelity to the LLI model and to gauge the level of literacy instruction provided in these groups. To ensure the reliability of data, observers received a manual which provided definitions of terms, examples and explanations of target strategies, and a description of procedures for completing the instruments. Observers also received training on the instrument in a group session and monitoring by CREP researchers throughout the observations.

LLI Online Data Management System Intervention Record

The LLI Online Data Management System (ODMS) is a tool developed by Heinemann to allow teachers to enter and track data for their LLI groups and individual students, including demographic information, entry and exit benchmark scores, Weekly Reading Record scores, attendance, lessons completed, and current reading level. This data management tool allows teachers and administrators to create individual, group, or school-level reports to monitor students' progress. The Intervention Record in the ODMS was used for tracking student and teacher attendance, reasons for absence, student reading selections, and achievement levels. When possible, CREP utilized teacher-provided intervention records from ODMS to provide an additional measure of the LLI implementation fidelity at each school, particularly with regard to the 30-minutes-a-day, 5-days-a-week instructional cycle.

Quality of Core Literacy Instruction

Literacy Observation Tool (LOT)

The Literacy Observation Tool (LOT) was developed by researchers at CREP to serve as an instrument for observing in elementary classrooms where teachers are engaged in teaching reading and other literacy-related practices. The LOT has been aligned to the National Reading Panel and National Research Council findings. It captures explicit instruction in the five essential components of reading identified by the National Reading Panel as important in achieving effective reading instruction: Phonemic Awareness, Phonics, Fluency, Text Comprehension, and Vocabulary. Standard use of the LOT involves multiple classroom observations during a designated literacy block (typically 1.5 to 2 hours), with seven to nine classrooms each observed for 10 minutes. In a study of 70 schools across Tennessee, strong evidence was established for the reliability of the LOT, with a phi coefficient of .75 for five observations and .82 for eight observations at a school (Sterbinsky & Ross, 2003).

Twice during the 2015-2016 school year (once at the beginning and once at the end), site researchers trained by CREP conducted a set of seven to nine 10-minute LOT observations in the regular 3rd-5th grade classrooms at each participating school. Each set of observations was conducted in one day during the school's literacy block, and the ratings from the seven to nine individual classroom observations were combined to form a single LOT composite for that school. Therefore, the LOT was used to obtain a measure of the quality of the regular classroom literacy instruction received by students in the study by taking a "snapshot" of each school's core literacy instruction. To ensure that

the identifying and coding of literacy instructional variables occurred in a consistent manner, observers received formal training, user's manuals, and monitoring by CREP researchers.

School and Home Support for Literacy

Leveled Literacy Intervention Teacher Questionnaire – Revised (LLITQ-R)

The Leveled Literacy Intervention Teacher Questionnaire – Revised (LLITQ-R), developed by CREP researchers for a previous study of LLI, was used in this study as a measure of the participating LLI teachers' views of the efficacy of LLI, their implementation of the LLI model, and their students' progress in literacy, as well as the overall support for literacy and LLI in their schools. The LLITQ-R consists of 23 items on a 5-point scale ranging from 1 (Strongly Disagree) to 5 (Strongly Agree), five items on a 4-point scale ranging from 0 (Not at All) to 3 (Extensively), five items on a 5-point scale ranging from 0 (Not at All/Never) to 4 (Regularly/Every Day), and three open-ended items regarding LLI's strengths and areas for improvement as well as reasons to continue or not continue using the LLI system. The LLITQ-R was administered to participating LLI teachers at the beginning and end of the school year.

Classroom Teacher Literacy Instruction Questionnaire (CTLIQ)

The Classroom Teacher Literacy Instruction Questionnaire (CTLIQ), also developed by CREP for a previous study of LLI, was used in the current study to measure the overall support for literacy in the participating schools and the nature of the regular classroom literacy instruction received by the students in the study. The CTLIQ assessed 3rd-5th grade classroom teachers' self-reported literacy instructional practices and their perceptions of the core literacy program at their schools, as well as their perceptions of LLI. The CTLIQ consists of 24 items on a 5-point scale ranging from 1 (Strongly Disagree) to 5 (Strongly Agree), five items on a 4-point scale ranging from 0 (Not at All) to 3 (Extensively), 10 items on a 5-point scale ranging from 0 (Not at All/Never) to 4 (Regularly/Every Day), and three open-ended items regarding the core literacy program's strengths and areas for improvement as well as reasons to continue or not continue the program. The CTLIQ was administered to 3rd-5th grade classroom teachers at the participating schools at the beginning and end of the school year.

Leveled Literacy Intervention Principal Questionnaire (LLIPQ)

CREP researchers developed the Leveled Literacy Intervention Principal Questionnaire (LLIPQ) for the current study as a measure of school support for literacy at the administrative level as well as principals' support for LLI specifically. The LLIPQ assessed principals' perceptions of their schools' core literacy program, their understanding of and familiarity with the LLI system, and their perceptions of LLI's implementation and efficacy at their schools. The LLIPQ consists of 26 items on a 5-point scale ranging from 0 (Strongly Disagree) to 4 (Strongly Agree) and three items on a 4-point scale ranging from 0 (Not at All) to 3 (Extensively). In addition, the survey contains five open-ended items designed to ascertain principals' perceptions of the strengths and areas for improvement of the LLI system, reasons to continue or not continue using the system, challenges to LLI implementation, additional resources needed, and efficacy of LLI in comparison with other supplemental literacy interventions at the principals' schools. The LLIPQ was administered to principals of the schools participating in the study at the end of the school year.

Home Literacy Support Questionnaire (HLSQ)

In order to measure the amount of support for literacy received by participating students at home, CREP researchers developed the Home Literacy Support Questionnaire (HLSQ). The HLSQ is a brief survey that asks parents/guardians (or other caretaking family members) about literacy activities in which their child may engage at home, as well as their own involvement in and encouragement of these activities. Additionally, the HLSQ assesses respondents’ feedback on LLI, if applicable. The HLSQ is comprised of 14 items on a 5-point scale ranging from 0 (Strongly Disagree) to 4 (Strongly Agree) and two open-ended items regarding respondents’ perceptions of the strengths and areas for improvement of their child’s literacy instruction at school. The HLSQ was provided in both English and Spanish versions, and was administered to parents/guardians of both treatment and control group students in the study at the end of the school year.

Stakeholder Feedback

LLI Teacher and Site Researcher Focus Groups

Voluntary, structured focus groups were conducted with LLI teachers and site researchers at the end of the school year to provide more information about the overall climate for literacy instruction in the district as well as additional feedback on the LLI system. Participating LLI teachers responded to questions regarding LLI’s strengths and areas for improvement, the efficacy of LLI in meeting students’ needs, support within their schools for LLI, and their opinion of the LLI professional development they received. Finally, site researchers discussed their perceptions of LLI’s strengths and areas for improvement, students’ response to LLI, the quality of LLI and core literacy instruction received by students in the study, and their opinion of the training and use of the data collection instruments for the study.

Instrumentation Summary

Table 6 summarizes each of the research questions and the participants and provides the data sources and methodology used to investigate each question.

Table 6: Summary of Data Sources and Participants by Research Question

Research Questions	Participants	Data Sources	Method
1) What progress in literacy achievement, if any, do students who receive LLI in grades 3-5 make compared to students who receive core literacy instruction alone? a) Does the effectiveness of LLI vary by the following subgroups: English Language Learners, students with a special education designation, and ethnic minorities (i.e., African-American and Hispanic students)?	<ul style="list-style-type: none"> • LLI treatment and control students • LLI and classroom teachers • Principals 	<ul style="list-style-type: none"> • Fountas & Pinnell Benchmarks • State Assessments • LLI teacher survey (LLITQ-R) • LLI teacher focus group • Classroom teacher survey (CTLIQ) • Classroom literacy observations (LOT) • Principal survey (LLIPQ) 	<ul style="list-style-type: none"> • Quantitative assessments of student progress in reading achievement • Qualitative assessment of student progress through teacher and administrator feedback • Quantitative data regarding regular classroom literacy instruction

Table 6: Continued

Research Questions	Participants	Data Sources	Method
2) At what level of fidelity to the program model is LLI implemented by teachers participating in the study?	<ul style="list-style-type: none"> • LLI teachers • Site researchers 	<ul style="list-style-type: none"> • LLI observations (LLIOT) • LLI Online Data Management System Intervention Records • LLI teacher survey (LLITQ-R) • LLI teacher focus groups • Site researcher focus groups 	<ul style="list-style-type: none"> • Quantitative and qualitative assessments of LLI instructional strategies and delivery
3) What are stakeholders' perceptions of the LLI system for grades 3-5 and the core literacy program?	<ul style="list-style-type: none"> • LLI teachers • Classroom teachers • Parents/guardians • Principals 	<ul style="list-style-type: none"> • LLI teacher survey (LLITQ-R) • LLI teacher focus groups • Classroom teacher survey (CTLIQ) • Parent/guardian survey (HLSQ) • Principal survey (LLIPQ) 	<ul style="list-style-type: none"> • Quantitative and qualitative assessment of LLI teachers' perceptions regarding LLI's impact on their instruction and their students' literacy, as well as classroom teachers', parents'/guardians', and principals' perceptions of LLI and the core literacy program in general

Procedure

The current study extended from Fall 2014 through Fall 2016. In Fall 2014 and Spring 2015, CREP researchers worked with the sponsor to develop a research plan and select school districts as sites for the research. Three school districts, including SPS, agreed to participate and were chosen due to varying regional locations and student populations (i.e., a high percentage of ELL, minority, and economically disadvantaged students), and the established relationships with the sponsor and with LLI. Although the original study proposal included randomization of eligible students to treatment and control groups, this methodology was replaced with a quasi-experimental matched-pair design to better serve districts' needs. Participating districts agreed to deliver LLI as designed by the developers, allow the Fountas & Pinnell Benchmark Assessments to be administered to students in the study, and provide the researchers with individual student-level data (e.g., demographic information, district-selected state assessment scores).

In May and June 2015, CREP researchers conducted site visits in Sandwich, MA to meet with key district-level administrators as well as principals and teachers, when possible, at schools interested in participating in the study. The research team provided an overview of the study requirements and the incentives to participate, which included – for each school – a set of Fountas & Pinnell Benchmark Assessment Systems 1 and 2 (as needed), Heinemann professional development on the BAS and LLI, as well as a \$1500 per semester stipend, complimentary use of the LLI Online Data Management System, and a school-level report of results as requested by school principals. One school in Sandwich, MA participated in the current study.

During Summer 2015, CREP researchers refined the existing data collection instruments, which were developed for a previous LLI study in 2011, obtained IRB approval for the study, and worked with each district's staff to identify school coordinators and site researchers to conduct ongoing local work for the study. One school coordinator was identified from the team of LLI teachers and instructional coordinators at each participating school to coordinate data collection activities with CREP and help

ensure smooth LLI implementation. Additionally, site researcher applicants were selected from pools of local-area educators, primarily retired teachers in each district; these applicants formed teams of site researchers to collect data for the study (i.e., Fountas & Pinnell benchmarks, LOT and LLIOT observations) throughout the school year. In Sandwich, two applicants were selected as site-researchers and collected data full-time for the study. Finally, a District Liaison, with an administrative position in the district, was selected in Sandwich to assist with district-level data and overall coordination of the project.

During early Fall 2015, CREP researchers returned to each district to meet with district personnel, site researchers, and school coordinators in order to finalize the study timelines and logistics, including plans for pretesting students in the study on the Fountas & Pinnell benchmarks between late August and early October. The CREP research team also worked with Heinemann consultants and the district professional development teams to provide training to the site researchers and LLI teachers who would be participating in the study, which included the following: (1) two full days of LLI training for teachers and site researchers (a follow-up training day was conducted in each district according to district convenience); (2) a full-day training on the Fountas & Pinnell Benchmark Assessment System for LLI teachers and site researchers. Training for site researchers on the LLI and classroom observation tools (i.e., LLIOT and LOT) also occurred each day following the Heinemann-led training sessions. One Heinemann consultant conducted the LLI and benchmark trainings, while two CREP researchers conducted the observation trainings. In early Spring 2016, CREP researchers led refresher trainings for site researchers on the Fountas & Pinnell benchmarks and the LOT and LLIOT observations. These refresher trainings occurred prior to the second round of LLIOT and LOT observations and the benchmark post-testing.

At the beginning of the 2015-2016 school year, parental consent forms were distributed in each district to students in grades 3-5 who were preliminarily identified by classroom teachers as students who might qualify for the LLI study. The CREP research team also worked with the district coordinators/liaisons and school coordinators to develop lists of students in grades 3-5 who were eligible to participate in the study. Selection criteria included students who would be able to receive delayed literacy intervention services if assigned to the control group, could receive instruction in English, were not known to demonstrate high absenteeism, were below grade level based on each district's state or district testing standards according to scores in Spring 2015. CREP researchers also worked with the district coordinator and school coordinators to obtain active consent from LLI and classroom teachers who would be taking part in the study.

Once eligible students were identified and parental consent was received, pretesting of these students with the Fountas & Pinnell benchmarks was conducted by the site researchers. Subsequently, CREP researchers and statisticians conducted a matched-pairs analysis to match students based on demographic characteristics (i.e., gender, ethnicity, ELL status, special education, and economically disadvantaged status), and Spring 2015 Fountas & Pinnell benchmark scores of instructional reading level. Students in the treatment groups were then placed in LLI groups by LLI teachers, and the planned 24 weeks of LLI instruction for students began. The starting date for LLI varied across the participating districts and schools due to varying school-year academic calendars, the length of time needed to identify eligible students, obtain consent, complete matching, administer the benchmarks, and organize and schedule the LLI groups; however, all groups began LLI by October 2015. Posttests with the Fountas & Pinnell benchmarks for all students were completed in Spring 2016.

Consistent with LLI program recommendations, the LLI research period lasted for a minimum of 24 calendar weeks, excluding the two weeks that districts were out of session for the winter holidays and other holidays in each district. During this 24-week period, control group students did not receive LLI; however, they could receive it after the research period if they still needed it (e.g., according to teacher judgment or post-benchmark scores). Site researchers used the LLIOT to conduct two random observations of each LLI group, one in Fall 2015 and one in Spring 2016; additionally, they conducted the first set of LOT (classroom literacy) observations in Fall 2015. A series of partnered observations using the LLIOT were also conducted in groups (five site researchers and one CREP researcher) during the round of observations in Fall 2015, in order to allow an assessment of inter-rater reliability.

In Fall 2015 and Spring 2016, the school and district coordinators/district liaisons were asked to encourage all 3rd-5th grade LLI and classroom teachers with students in the study to complete the online LLI Teacher Questionnaire – Revised (LLITQ-R) or Classroom Teacher Literacy Instruction Questionnaire (CTLIQ), as applicable. Additionally, the principals at the participating schools were asked to complete the online LLI Principal Questionnaire (LLIPQ) at the end of the school year. CREP assisted in the online survey process by providing instructions and log-in information to all participants. CREP also distributed paper copies of the Home Literacy Support Questionnaire (HLSQ) to the school coordinators, who sent them home with both treatment and control group students. Parents/guardians could complete the survey and return it to the child’s school, where it was collected by the school coordinator, or return it directly to CREP via mail.

The CREP research team held end-of-year meetings with each district, including school coordinators, district coordinators/liaisons, participating LLI teachers, and site researchers to debrief them, discuss any remaining issues, and conduct structured focus groups. The purpose of the focus groups was to collect qualitative data related to the study, the LLI materials, and participants’ individual and collective views of LLI. CREP researchers also met individually with LLI teachers, as needed, to verify teachers’ group and student data taken from the LLI Online Data Management System and to request missing information. Finally, demographic and state assessment achievement data (i.e., Spring 2015 and Spring 2016 state assessment scores) were provided electronically for students by each district in Summer and Fall 2016.

Table 7 provides a summary of data collection procedures, including the instruments organized by type, a general timeline and description of the data collection process, and the number received for each instrument.

Table 7: Data Collection Summary

Type of Measure	Instrument	Timeline	Description
Student Achievement Measures	<ul style="list-style-type: none"> Fountas & Pinnell Benchmarks State Assessments (Literacy) 	<ul style="list-style-type: none"> September/October 2015 April - June 2016 Spring 2015 Spring 2016 	<ul style="list-style-type: none"> The Fountas & Pinnell Benchmarks were administered to all students in both treatment and control groups as a pretest in Fall 2015 and a posttest in Spring 2016. Both scale and proficiency scores in literacy were collected from each district for participating students in grades 4 and 5 for Spring 2015 and Spring 2016, for whom scores were available.

Table 7: Continued

Type of Measure	Instrument	Timeline	Description
Observations	<ul style="list-style-type: none"> • LLIOT • LOT 	<ul style="list-style-type: none"> • September/October 2016 • March-May 2016 • October/November 2016 • March-May 2016 	<ul style="list-style-type: none"> • Trained on-site researchers observed all 3rd-5th grade LLI groups twice--once in Fall 2015 and once in Spring 2016. • Trained on-site researchers observed randomly selected 3rd-5th grade classrooms during the literacy block at the beginning and end of the school year.
Surveys	<ul style="list-style-type: none"> • LLITQ • CTLIQ • LLIPQ • HLSQ 	<ul style="list-style-type: none"> • September/October 2015 & May/June 2016 • September/October 2015 & May/June 2016 • May/June 2016 • May/June 2016 	<ul style="list-style-type: none"> • Surveys were completed at the beginning and end of the school year to obtain feedback from LLI teachers and 3rd-5th grade classroom teachers. • Surveys were completed at the end of the school year to obtain feedback from principals and parents/guardians.
Focus Groups	<ul style="list-style-type: none"> • LLI Teachers • Site Researchers 	<ul style="list-style-type: none"> • September/October 2015 & May/June 2016 • September/October 2015 & May/June 2016 	<ul style="list-style-type: none"> • Focus groups were completed at the beginning and end of the school year to obtain qualitative feedback about LLI and students' progress from LLI teachers. • Focus Groups were completed at the beginning and end of the school year to obtain qualitative feedback about LLI and students' progress from site researchers.

LLI Dosage: Number of Days of LLI Instruction

Across all participating schools, the number of LLI instructional days for LLI students from student attendance data was available for nearly all of the treatment group (N = 37). For the students with LLI instructional days provided, treatment group students received, on average, 67 days of LLI instruction over the minimum of 24 calendar weeks between October 2015 and May 2016, with a range of 45 to 87 days of instruction.

The recommended amount of LLI instruction for 3rd-5th grade students, according to the LLI program guide, is 18-24 weeks (approximately 90-120 days). The average of 67 days of LLI instruction provided did not meet these recommended amounts. It is also important to note from the student attendance data provided that none of the treatment students received the recommended minimum of 90 days due to student-level factors (e.g., individual absences or unavailability during LLI group time) as well as school or district limitations (e.g., holidays, assessment windows during which LLI teachers and/or students were pulled during LLI group time, delays in starting LLI due to scheduling conflicts or difficulty accessing student data).

Results

The following section presents the results of the study, discussed in relation to each instrument and each grade level, as appropriate. First, a summary of the quantitative and qualitative results will be presented, and the conclusion section will further discuss these results as they pertain to each of the research questions in the present study.

Student Literacy Achievement

To determine whether LLI students' progress in literacy was statistically significantly or substantively different from that of their control group counterparts, analyses of pretest to posttest gains were conducted at Grades 3-5 with respect to three measures of literacy achievement: the Fountas & Pinnell Benchmark Assessment System (BAS), state Literacy Assessment proficiency levels, and state Literacy Assessment scale scores.

Benchmark Assessment System (BAS)

At all three grades, BAS gains were computed by obtaining a numeric equivalent for students' beginning and ending instructional levels (e.g., A = 1, B = 2, and so on) and then computing the difference between the posttest and pretest numeric scores. Students who scored below the lowest level, A, were assigned a level of "Pre-A" with a corresponding score of 0.

State assessment proficiency level

Gains on the state Literacy Assessments were computed based on students' proficiency level, recoded as a dichotomous variable (Proficient = 1, Not Proficient = 0) for the Spring 2015 administration (pretest) and Spring 2016 administration (posttest). In a manner similar to computing benchmark gains, Proficiency Level gains were derived by subtracting posttest from pretest "levels." However, students who were in the Proficient level for both Spring 2015 and Spring 2016 also received a gain score of one (1) as that was considered to be a positive outcome. While benchmark gain scores have a wider range of 27 (0 through 26), Proficiency Level gain scores have a much smaller range of only 3 (-1 through 1).

State assessment scale scores

In addition to the proficiency level, gains on the state literacy assessments were also computed on students' scale scores, which were derived by directly subtracting posttest from pretest scale scores. For the three measures of literacy gains, differences between LLI (treatment) and control groups overall, and by various subgroups based on students' demographic characteristics, such as ethnicity, English Language Learner (ELL) status, Special Education status, and Economically Disadvantaged status, were compared within grade level via multiple analysis of covariance (ANCOVA) procedures, with no correction applied to the probability level of the statistical outcomes, as these analyses were regarded as exploratory. In addition, differences in gains at each grade level were explored by classifying students' based on their achievement level on the Fall 2015 benchmark as either (1) "Low-Achievers" (scoring **at or below** the median for the combined (i.e., both LLI and control) study sample on the Fall 2015 benchmark assessment) or (2) "High-Achievers" (scoring **above** the median for the combined study sample on the Fall 2015 benchmark assessment). Although both groups would be considered low achieving based on the fact that they were receiving LLI, the "Low-Achievers" group closely relates to the benchmark level cut score for "Does Not Meet Expectations: Needs Intensive Intervention" (i.e., this group would be considered the "lowest of the Low-Achievers" so to speak). The mean gains for these groups were also statistically tested using ANCOVA procedures, with no corrections for multiple comparisons. The covariates in the ANCOVA models included dummy variables for Minority (i.e., non-White), ELL status, Special Education status, and Economically Disadvantaged status. A total of eight comparisons were made: All students, White, African-American, Hispanic, Minority, Special Education, Limited English Proficient, and Economically Disadvantaged.

In addition to testing for statistical significance, an effect size was calculated. As an indicator of the impact or “practical significance” of the treatment, the “effect size” (calculated as Hedges’ *g*) is a descriptive statistic that indicates the magnitude of the difference (in standard deviation units) between two measures. For example, a positive effect size would indicate a larger (i.e., better) LLI group gain, while a negative effect size would indicate a larger (i.e., better) control group gain. Based on the guidelines from the What Works Clearinghouse (WWC), a unit within the research division of the U.S. Department of Education, an effect size of +/- 0.25 is considered to be “substantively important” (i.e., educationally meaningful) (What Works Clearinghouse, 2014).

Three types of ANCOVA analyses were conducted for the current report, namely:

(1) when LLI and control group students had baseline equivalence (i.e., neither group had an advantage on the pretest) and LLI students demonstrated either statistically significantly and/or substantively larger gains compared to their control counterparts with respect to any of the three outcome measures;

(2) LLI students showed larger gains in outcomes despite control group students having a substantively important advantage at baseline; and

(3) when control group students had a substantively important baseline advantage and an advantage on the outcome, but the outcome difference favoring the control group was not substantively important.

(1) Baseline equivalence + LLI group outcome advantage

According to the ANCOVA analyses, **no** significant findings for any group in Sandwich were demonstrated with this type of result.

(2) Control group baseline advantage + LLI group outcome advantage

For some subgroups, LLI students showed larger gains despite control group students having a substantively important advantage at the baseline (see Table 8). Here we define “larger” as either substantively important and/or statistically significant, or simply larger gains, even though neither statistically significant nor substantively important.

For students at Sandwich, LLI students in the following subgroups in 3rd grade and 4th grade exhibited substantively larger benchmark level (BAS) gains compared to their control counterparts:

- 3rd grade aggregate [All Students] ($g = 0.44$);
- 3rd grade White students ($g = 0.43$);
- 4th grade aggregate [All Students] ($g = 0.65$);
- 4th grade White students ($g = 0.64$);
- 4th grade High-Achiever students ($g = 0.64$);
- 4th grade High-Achiever White students ($g = 0.64$);

In addition, LLI students in two subgroups in both 3rd and 5th grade showed larger gains on benchmark levels (BAS) compared to their control counterparts, even though the difference was neither statistically significant nor substantively important.

- 3rd grade High-Achiever students ($g = 0.21$);
- 3rd grade High-Achiever White students ($g = 0.21$);
- 5th grade aggregate [All Students] ($g = 0.16$);
- 5th grade White students ($g = 0.16$);

It should be noted that all 5th grade students in the sample are White; in other words, the 5th grade White subgroup is the same as the 5th grade aggregate [All Students] group.

Table 8: Control and LLI Gains: Control Students with Baseline Advantage/LLI Students with Outcome Advantage

Grade level	Student group	Control			LLI			F	p	g
		n	M	SD	n	M	SD			
Fountas & Pinnell Benchmarks										
3	All	26	1.95	1.16	8	2.54	1.77	1.145	0.293	0.44
	White	25	1.95	1.19	8	2.54	1.77	1.145	0.293	0.43
	High-Achiever	23	1.96	1.22	4	2.21	0.50	0.144	0.708	0.21
	High-Achiever White	22	1.96	1.25	4	2.21	0.50	0.144	0.708	0.21
4	All	24	1.70	1.65	16	2.83	1.83	3.584	0.067	0.65
	White	23	1.71	1.68	16	2.85	1.83	3.584	0.067	0.64
	High-Achiever	24	1.74	1.65	15	2.88	1.84	3.455	0.072	0.64
	High-Achiever White	23	1.76	1.68	15	2.90	1.84	3.455	0.072	0.64
5	All	17	1.78	1.36	14	1.98	1.14	0.186	0.670	0.16
	White	17	1.78	1.36	14	1.98	1.14	0.186	0.670	0.16

Note. Cells shaded in gray have very small sample sizes, and results should be treated with caution. Effect sizes in red are substantively important (i.e., ≥ 0.25)

(3) Control group baseline advantage + control group outcome advantage

In addition to the two types of results reported above, there were also outcomes where control students had a substantively important advantage at baseline, and also achieved larger gains than LLI students on the outcome, but the gain was not substantively important. This type of result would suggest that LLI students moved to close the achievement gap relative to their control counterparts, even though LLI students did not show better gains. However, according to the ANCOVA analyses, no subgroup in Sandwich demonstrated this type of result.

Summary

In looking at the three outcome measures, positive findings for LLI only emerged on the Benchmark. In all cases, the positive findings were cases where the control group had a substantively important advantage on the pretest, but LLI students demonstrated an advantage on the outcome.

- For the **Benchmark Assessment (BAS)**
 - The only positive findings were for the combined samples at all three grade levels, and High-Achievers in 3rd and 4th grades.

- In addition, only two groups demonstrated positive outcomes: All students and White students.
- Across the three outcomes,
 - All of the positive findings had small sample sizes, and in nearly all cases for just the LLI group, which limits the generalizability of the findings.
 - There were no positive findings for the Low-Achiever group at any grade level

However, it is important to note that the student samples in Sandwich for each grade level were very small overall and for all available subgroups. Therefore, the results of the current study may not be representative of the impacts of LLI on demographic subgroups of students (e.g., special education, minority, economically disadvantaged) in the district.

Intervention Fidelity

Leveled Literacy Intervention Observation Tool (LLIOT)

Descriptive Results

The Leveled Literacy Intervention Observation Tool (LLIOT) involved a targeted, 30-minute observation of LLI implementation and instructional strategies ($n = 23$ observations). Table 9 illustrates the frequencies for each item on the LLIOT, as observed during the visits. The results from the LLIOT revealed that 3 of the 19 components were rated “Acceptable” or “Excellent” at least 65.0% of the time. The highest rated lesson components (i.e., demonstrating the highest degree of implementation fidelity) included phonics and word study (69.6%), discussing and revisiting the text with reflections on the book (69.6%), and reading a new book with guided questions (69.5%). However, 8 of the 19 lesson components, revisiting a new book with a vocabulary focus, use of the prompting guide, and those related to test preparation, were not seen in any observations. The lesson components that needed the most improvement were reading a new book with students reading silently (13%) and reading a new book with assistance for students who need help (13%). Teachers were also rated highly (i.e., “Acceptable” or “Excellent”) on their use of literacy instructional strategies, including emphasizing understanding of what is read (82.6%) and engaging students in conversation about the text (82.6%). Further, in the majority of observed lessons materials needed were readily available (91.3%), students were actively engaged (82.6%), the teacher engaged in ongoing assessment of learning (78.3%), and the lesson was well organized (65.2%). Overall, observers perceived that the lesson was delivered as designed 34.8% of the time, and the average rating across all subscales of the LLIOT was 2.41 (i.e., between “Acceptable” [2] and “Excellent” [3]). All items can be found in Table 9 below.

Table 9: LLIOT Response Frequencies (n = 23)

Item	Percent Responded			
	Excellent	Acceptable	Needs Improvement	Not Observed
Quality of LLI Implementation				
Revisiting yesterday's new book- Comprehension focus	8.7	43.5	4.3	43.5
Revisiting yesterday's new book- Vocabulary focus	0.0	0.0	0.0	100.0
Revisiting yesterday's new book- Fluency focus	8.7	4.3	0.0	87.0
Phonics/word study (e.g., vowel sounds, suffixes, plurality, etc.)	17.4	52.2	4.3	26.1
New book - Guiding questions	21.7	47.8	13.0	17.4
New book - Students read silently	4.3	56.5	13.0	26.1
New book – Assist students who need help	0.0	56.5	13.0	30.4
Discussing and revisiting the text	8.7	60.9	8.7	21.7
Rereading and assessment- Setting specific purposes	17.4	17.4	8.7	56.5
Rereading and assessment- Listens to one student read	21.7	21.7	0.0	56.5
Shared and independent writing	0.0	47.8	4.3	47.8
Classroom and homework	0.0	13.0	4.3	82.6
Use of prompting guide	0.0	0.0	0.0	100.0
Test preparation - Think together - Explain test	0.0	0.0	0.0	100.0
Quality of LLI Implementation				
Test preparation – Think together – Recognize answer based on question	0.0	0.0	0.0	100.0
Test preparation- Have a try- Students read text	0.0	0.0	0.0	100.0
Test preparation- Have a try- Students identify main words	0.0	0.0	0.0	100.0
Test preparation- Have a try- Students organize thinking	0.0	0.0	0.0	100.0
Test preparation- On your own- Students read passage independently	0.0	0.0	0.0	100.0
Literacy Instructional Strategies				
Teacher models, encourages, and provides opportunities for fluent oral reading.	8.7	34.8	17.4	39.1
Teacher introduces vocabulary words (e.g., high frequency, story-specific words).	4.3	26.1	30.4	39.1
Teacher emphasizes understanding/comprehension of what is read.	26.1	56.5	8.7	8.7
Teacher models and encourages students to use appropriate reading strategies (e.g., phonemic awareness).	13.0	34.8	17.4	34.8
Literacy Instructional Strategies				
Teacher engages students in conversation about the text.	34.8	47.8	17.4	0.0
Teacher assists students in problem-solving.	17.4	43.5	17.4	21.7
Learning Environment				
Lesson is well organized.	34.8	30.4	34.8	0.0
Teacher appropriately paces lesson components.	8.7	30.4	56.5	4.3
Teacher engages in ongoing assessment of student learning (e.g., questioning, providing feedback/corrective instruction, checking responses).	17.4	60.9	13.0	8.7
Students are actively engaged.	65.2	17.4	17.4	0.0
Instructional modifications are observed when needed.	8.7	8.7	4.3	78.3
Instructional materials needed to implement lesson are readily available.	34.8	6.5	8.7	0.0
The lesson is delivered as designed.	0.0	34.8	60.9	4.3

The LLIOT also included items designed to describe the groups observed, which are summarized in Table 10. Results from these items indicated that observers most frequently saw 4th and 5th grade groups (30.4% and 43.4%, respectively). The majority of observed groups took place in a designated intervention area (95.7%) and all lasted between 35 and 50 minutes, which was consistent with LLI’s design. Further, most of the observed groups had four students, with the remaining groups having three or two. Finally, a greater number of even- than odd-numbered lessons were observed (52.2% and 34.8%, respectively). Novel study was also observed part of the time (13%). All items can be found in Table 10 below.

Table 10: LLIOT Summary Items (n = 23)

Item	Percent Responded
Grade Level	
3	26.1
4	30.4
5	43.5
Location of Group	
Intervention Area	95.7
Classroom	0.0
Other	4.3
Number of Students in Group	
1	0.0
2	13.0
3	17.4
4	69.6
5 or more	0.0
Total Instructional Minutes	
Less than 35	0.0
35 – 50	100.0
More than 50	0.0
LLI Lesson Number	
Even	52.2
Odd	34.8
Novel Study	13.0
Test Preparation	0.0

Note. Item percentages may not total 100% due to missing input or multiple responses from some participants.

Observers conducting the LLIOT also recorded open-ended comments summarizing the instructional materials used during the lesson and their perceptions of the quality of instruction, level of student participation, and overall success of the lesson. Observers’ comments were summarized using a structured, multi-step process. First, the original comments were assigned codes representing their basic content. Next, these codes were grouped into categories, which were then organized into overarching themes. Final analysis produced frequency percentages for each theme. Because it was possible for some comments to contain multiple content codes, the percentages reported reflect the total number of codes within each theme and not necessarily the total number of comments received from observers. Observers’ responses are summarized below by Fall 2015 and Spring 2016.

Fall 2015

In the Fall of 2015 for SPS, 45.5% of the comments were related to lesson resources and materials, 27.3% were related to the success of the lesson, 18.2% were related to quality of literacy instruction, and 9.1% were related to student participation and engagement. The majority of these comments (72.7%) were neutral or descriptive in nature, while 27.3% were positive. No comments indicated areas needing improvement. In general, observers reported that the lessons were appropriately paced, well-organized, and delivered according to LLI guidelines. Comments were made that acknowledged teachers made effective use of such strategies as monitoring, questioning, prompts, and reinforcement; the students were actively engaged, motivated to learn, and enthusiastic, and a wide variety of instructional materials were readily available. However, in general, the observers simply described the lesson materials without commenting on specific examples of teacher or student behavior. Sample comments from the observers are provided below.

"Purple system. Working long vowels- word cards (10 mins.) New book- asked guiding questions- intro'd voc words- pgs. (10 mins.) Silent read- invited to read- sound out syllables. Sticky note writing? Questioning- shared writer's message."

"1:45- Teacher gives overview- introduces Reader's Theatre and assigns parts. 1:56- Reader's Theatre- focus on fluency along w/ accuracy/ teacher models- even mistakes. 2:05- story reflections-> students offer different endings to the story. Teacher makes connections between story and real life situation. life teaches us lessons. Teachers [sic] says: 'Tell me about that...' -teacher doesn't give students answers, asks 'deep questions.' Teacher uses characters from the story to guide students. 2:20- teacher moves into writing component. Teacher clearly outlines assignment and provides graphic organizer to order thoughts. Students use graphic organizer to write non-fiction paragraphs. Students write until end of class."

Spring 2016

In the Spring of 2016 for SPS, 44.4% of the comments were related to the success of the lesson, 22.2% were related to lesson resources and materials, 22.2% were related to student participation and engagement, and 11.1% were related to quality of literacy instruction. The majority of these comments (55.6%) were positive in nature, while 22.2% were neutral or descriptive. Additionally, 22.2% of comments were related to areas needing improvement. In general, observers reported that the lessons were appropriately paced, well-organized, and delivered according to LLI guidelines. Comments were made that acknowledged teachers were well-organized, managed time well, and used materials effectively; the students were actively engaged, motivated to learn, and enthusiastic, and a wide variety of instructional materials were readily available. However, some observers noted that students were allowed to engage in off-topic discussion and sometimes were uncooperative. Sample comments from the observers are provided below.

"Observed running record & questions. Writing-- opinions on character's reasoning. Word Study-- syllables with /air/ air, aire, ear, are. Well organized-- students very cooperative."

"Diagraphs ph, sh, ch, th, wh-- Review (white boards). (Teacher) is very well organized-- Students read silently with a question to answer while (teacher) did running record. (was

able to listen in on running record) Shared Writing (20 mins) Didn't intro new book-- mentioned it as Monday's book.

“Had a difficult time observing (teacher). She was finishing up yesterday's even lesson, by doing Word Study, then she introduced a new book. They spent 35 mins on the new book., reading aloud/ silently/ writing about/ discussing. (Teacher) did all of these areas well, but I don't think she was following LLI protocols. Two members of her group never come on Monday because they have a 2nd Special that day. Makes things quite difficult for her pacing.”

Consistency of LLI Implementation

The LLIOT was conducted at both the beginning and end of each LLI group containing at least one treatment group student in order to measure any changes in implementation over time. For third through fifth grades, pretest observations were conducted in Fall 2015, and posttest observations were conducted in Spring 2016. The 24 individual LLIOT items were divided into and analyzed as three subscales: Quality of LLI Implementation (Items 1-11), Literacy Instructional Strategies (Items 12-17), and Learning Environment (Items 18-24), with each item rated on a three-point scale: Needs Improvement (1), Acceptable (2), and Excellent (3). For each subscale, a mean (i.e., average) across all of the items was calculated, and the means between the two time points tested via an independent *t*-test, as it was not possible to link individual observations across the two time points.

In addition to testing for statistical significance, an effect size was calculated. As an indicator of the impact or “practical significance” of the treatment, the effect size (calculated as Hedges’ *g*) is a descriptive statistic that indicates the magnitude of the difference (in standard deviation units) between two measures. For example, a positive effect size would indicate a higher (i.e., better) Spring 2016 mean, while a negative effect size would indicate a higher (i.e., better) Fall 2015 mean. Based on guidelines from the What Works Clearinghouse (WWC), a unit within the research division of the U.S. Department of Education, an effect size of +/- 0.25 is considered to be “substantively important” (What Works Clearinghouse, 2014). Results are summarized by grade level below.

Overall, it appears that LLI instruction remained consistently acceptable or decreased slightly throughout the program across all three grade levels, with substantively important improvement in quality of implementation for fifth grade and substantively important decreases in literacy instruction strategies for fifth grade, and learning environment for fifth grade. It should be noted that for third grade, all subscales were rated between needing improvement and acceptable at both time points.

3rd Grade

The three independent *t*-tests that contrasted teacher behaviors at times one and two conducted on the set of means obtained on the LLIOT’s ten-item “Quality of LLI Implementation” scale ($t = 0.031, p = 0.977, g = -0.02$), the six-item “Literacy Instructional Strategies” scale ($t = 0.256, p = 0.811, g = -0.22$), and the seven-item “Learning Environment” scale ($t = 0.122, p = 0.909, g = -0.10$), showed no statistically significant or substantively important differences observed for any of the two comparisons. The average rating was between “Need Improvement” (1.00) and “Acceptable” (2.00) for each subscale at both time points (see Table 11).

4th Grade

The descriptive statistics and independent *t*-test results for each of the three LLIOT subscales for the fourth grade groups are presented in Table 11. There were no statistically significant or substantively important differences between the pretest and posttest observations for the three subscales: “Quality of LLI Implementation” ($t = -0.204, p = 0.846, g = 0.15$), “Literacy Instructional Strategies” ($t = -0.274, p = 0.795, g = 0.21$), and “Learning Environment” ($t = 0.082, p = 0.938, g = -0.07$). For all subscales, the average rating was between “Acceptable” (2.00) and “Excellent” (3.00) at both time points.

5th Grade

The results for the three independent *t*-tests for the fifth grade groups observed at the posttest revealed no statistically significant differences between the two sets of observations for the each of the subscales. However, there was substantively important improvement from pretest to posttest for “Quality of LLI Implementation” ($t = -0.465, p = 0.654, g = 0.30$), while there was substantively important decreases from pretest to posttest for “Literacy Instructional Strategies” ($t = 0.717, p = 0.494, g = -0.45$) and “Learning Environment” ($t = 1.286, p = 0.234, g = -0.82$). The average rating was between “Acceptable” (2.00) and “Excellent” (3.00) for each subscale at both time points, with the exception of the posttest mean for learning environment, which was between “Needs Improvement” (1.00) and “Acceptable” (2.00). Descriptive statistics and independent *t*-test results are summarized in Table 11 below.

Table 11: Independent T-Test Results for LLIOT Subscales by Grade Level

Achievement Measure	Pretest			Posttest			<i>t</i>	<i>p</i>	<i>g</i>
	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>			
Quality of LLI Implementation									
3rd Grade	3	1.76	0.08	3	1.75	0.66	0.031	0.977	-0.02
4th Grade	4	2.17	0.33	3	2.21	0.26	-0.204	0.846	0.15
5th Grade	5	2.19	0.33	5	2.26	0.09	-0.465	0.654	0.30 [^]
Literacy Instructional Strategies									
3rd Grade	3	1.40	0.35	3	1.33	0.29	0.256	0.811	-0.22
4th Grade	4	2.24	0.53	3	2.33	0.08	-0.274	0.795	0.21
5th Grade	5	2.22	0.48	5	2.00	0.47	0.717	0.494	-0.45 [^]
Learning Environment									
3rd Grade	3	1.41	0.22	3	1.39	0.23	0.122	0.909	-0.10
4th Grade	4	2.22	0.46	3	2.19	0.34	0.082	0.938	-0.07
5th Grade	5	2.28	0.34	5	1.97	0.40	1.286	0.234	-0.82 [^]

* $p < .05$

[^] Substantively important effect size (i.e., $g \geq 0.25$)

Quality of Core Literacy Instruction

Literacy Observation Tool (LOT)

Descriptive Results

The Literacy Observation Tool (LOT) involved seven to nine 10-minute observations of core literacy instruction in grades 3-5 during each school semester throughout the course of the study ($n = 2$ LOTs in Fall 2015 and $n = 2$ LOTs in Spring 2016). The LOT is designed to capture explicit instruction in the five essential components of reading: Phonemic Awareness, Phonics, Fluency, Text Comprehension, and Vocabulary. Whole group instruction was seen frequently or extensively during both observation time points. Learning environments that were conducive to cooperative interactions and that actively engaged students were also frequently or extensively observed during both time points, as were effective classroom management and teacher monitoring. Classroom libraries were also frequently noted by CREP observers during both semesters, as was evidence of student writing products. However, although some of the literacy activities in the LOT are more frequently a part of literacy instruction in grades K-2 rather than grades 3-5, it is important to note that some activities within all five of the essential reading components were rarely observed or not observed at all during the Fall 2015 and Spring 2016 observations, such as small group instruction, learning centers, discussion of print conventions, phonemic awareness instruction, joint oral reading, explicit instruction of comprehension strategies, writing instruction, and some forms of assessment. Also, some materials were rarely or never used, including basal texts, audio books, computers, newspapers, and worksheets. Table 12 illustrates the frequencies for each item on the LOT, as observed during the visits.

Table 12: LOT Response Frequencies

Literacy Observation Tool Data Summary (LOT) Items	% Rarely or Not Observed		% Occasionally		% Frequently or Extensively	
	Fall 2015	Spring 2016	Fall 2015	Spring 2016	Fall 2015	Spring 2016
Instructional Orientation						
Small group	0.0	100.0	0.0	0.0	100.0	0.0
Whole class	0.0	0.0	0.0	0.0	100.0	100.0
Learning centers	100.0	100.0	0.0	0.0	0.0	0.0
Cooperative/Collaborative learning	50.0	50.0	50.0	50.0	0.0	0.0
Concepts of Print						
Book/print conventions	100.0	100.0	0.0	0.0	0.0	0.0
Alphabetics						
Letter naming/knowledge	100.0	100.0	0.0	0.0	0.0	5.0
Phonemic awareness instruction	100.0	100.0	0.0	0.0	0.0	5.0
Rhyming	100.0	50.0	0.0	50.0	0.0	5.0
Explicit phonics instruction	50.0	50.0	0.0	0.0	50.0	50.0
Fluency						
Models fluent oral reading	50.0	50.0	0.0	0.0	50.0	50.0
Has student(s) read/reread orally (together)	50.0	100.0	50.0	0.0	0.0	0.0
Vocabulary						
Introduces/reviews key vocabulary	0.0	50.0	50.0	0.0	50.0	50.0
Explicit vocabulary instruction	50.0	50.0	0.0	0.0	50.0	50.0

Table 12: Continued

Literacy Observation Tool Data Summary (LOT) Items	% Rarely or Not Observed		% Occasionally		% Frequently or Extensively	
	Fall 2015	Spring 2016	Fall 2015	Spring 2016	Fall 2015	Spring 2016
Text Comprehension						
Explicit comprehension strategy instruction	50.0	100.0	0.0	0.0	50.0	0.0
Makes connection to prior knowledge	0.0	50.0	50.0	50.0	50.0	0.0
Asks students for predictions	0.0	50.0	50.0	0.0	50.0	50.0
Uses higher level questioning	50.0	50.0	0.0	0.0	50.0	50.0
Guides visual imaging	50.0	50.0	50.0	50.0	0.0	0.0
Guides interactive discussion	50.0	50.0	0.0	0.0	50.0	50.0
Independent Reading - The Student:						
Reads self-selected materials	0.0	50.0	0.0	0.0	100.0	50.0
Writing - The Teacher:						
Letter formation/handwriting	100.0	50.0	0.0	50.0	0.0	0.0
Writing process	50.0	50.0	0.0	0.0	50.0	50.0
Language mechanics lessons	50.0	50.0	0.0	50.0	50.0	0.0
Conference with students	50.0	100.0	0.0	0.0	50.0	0.0
Provides for students sharing	50.0	100.0	0.0	0.0	50.0	0.0
Writing - The Student:						
Writes independently	50.0	50.0	0.0	50.0	50.0	0.0
Response writing	0.0	0.0	0.0	50.0	100.0	50.0
Assessment						
Formal testing	100.0	100.0	0.0	0.0	0.0	0.0
Portfolios	100.0	50.0	0.0	50.0	0.0	0.0
IRI, running records	100.0	50.0	0.0	50.0	0.0	0.0
Learning Environment						
Conducive to cooperative interactions	0.0	0.0	0.0	50.0	100.0	50.0
Students actively engaged	0.0	0.0	0.0	0.0	100.0	100.0
Effective classroom management	0.0	0.0	0.0	0.0	100.0	100.0
Teacher actively monitors	50.0	0.0	0.0	0.0	50.0	100.0
Visible Print Environment						
Alphabet	0.0	50.0	50.0	50.0	50.0	0.0
Word wall	100.0	100.0	0.0	0.0	0.0	0.0
Labeling (names, objects, areas)	50.0	50.0	0.0	0.0	50.0	50.0
Classroom library	0.0	0.0	0.0	0.0	100.0	100.0
Evidence of student writing/work products	0.0	0.0	50.0	0.0	50.0	100.0
Materials Used						
Basal texts	100.0	100.0	0.0	0.0	0.0	0.0
Big books	100.0	100.0	0.0	0.0	0.0	0.0
Books on tape	100.0	100.0	0.0	0.0	0.0	0.0
Computers	100.0	100.0	0.0	0.0	0.0	0.0
Fiction books	0.0	0.0	0.0	50.0	100.0	50.0
Non-fiction books	0.0	0.0	50.0	50.0	50.0	50.0
Materials Used						
Poetry	50.0	0.0	50.0	50.0	0.0	50.0
Newspaper/magazines	100.0	100.0	0.0	0.0	0.0	0.0
Word/vocabulary materials	0.0	50.0	50.0	0.0	50.0	50.0
Worksheets/workbooks	0.0	100.0	100.0	0.0	0.0	0.0
Materials Used - Other	0.0	0.0	0.0	50.0	100.0	0.0

There were notable changes across observation time points that are worth addressing. For instance, use of additional materials, small group instruction, and fiction books decreased considerably from Fall 2015 to Spring 2016. Also, from Fall 2015 to Spring 2016, the use of poetry and presence of student written work increased. Some of these changes could be attributed to the heightened focus on academic achievement tests during the spring semester. Given the increased pressure on student and teacher performance in educational settings over the past several years, teachers may be utilizing reading activities that they feel are most beneficial to student achievement on standardized tests as the testing season approaches, particularly during the spring semester.

Given the small number of comments provided, site researchers' open-ended responses for Fall 2015 and Spring 2016 are summarized in general by question below. Observers conducting the LOT noted comments related to the strengths and areas for improvement across the observed classrooms during their school visits, as well as notes on the students' progress and recommendations for next steps.

When asked to describe the strengths of the classroom literacy programs observed, site researchers frequently mentioned comfortable, inviting classroom environments that were conducive to learning. In addition, observers listed student engagement as another fundamental component when discussing strengths. Site researchers observed that classroom materials and print-rich environments helped contribute to the high quality of classrooms. They reported that students were engaged in guided reading groups and independent reading and were receiving extensive word study opportunities. Site researchers also noted the use of literacy instruction as a strength, which seemed to focus on word work, vocabulary, and comprehension.

When asked to discuss concerns they had regarding the classroom literacy lessons that they observed, most site researchers' responses reflected the theme of general instructional strategies and classroom supports. In particular, some mentioned a lack of visual aids, such as the alphabet, while others mentioned not observing guided reading. However, in this case, observers noted that this might have been due to a school-wide focus on upcoming standardized assessments.

When asked to discuss student progress and next steps regarding the classrooms that they observed, site researchers mentioned that students were making progress. Site researchers commented that this progress could be attributed to the availability of instruction and independent work. They also mentioned student familiarity with all aspects of the program and engagement in interactive, though not collaborative, discussion could be helping students. Finally, observers stated that teachers were successfully taking advantage of instructional time but could focus on additional literacy strategies moving forward such as visual imagery and greater collaborative work.

Strengths

"Students actively engaged. Teachers comfortable w/ various areas of the literacy collaborative program."

Concerns

"I believe I did not observe guided reading taking place because teachers have begun benchmark assessments."

Progress

“Students seem to have available to them the instruction they need to progress.”

Next Steps

“Guiding visual imagery may need more emphasis. Instruction encourages students to interact but extensive collaboration was not observed. Perhaps this happens more frequently in other subject areas. Discussion however, is strongly interactive, but could be more so and probably will be as the year proceeds and students learn these strategies.”

School and Home Support for Literacy

Leveled Literacy Intervention Teacher Questionnaire – Revised (LLITQ-R)

The Leveled Literacy Intervention Teacher Questionnaire – Revised (LLITQ-R) was administered online to LLI teachers as a general measure of their implementation and perceptions of LLI, with five responses gathered in Fall 2015. Because only one response was recorded for Spring 2016, this semester was left out to protect confidentiality. Table 13 illustrates the frequencies of responses for each item on the LLITQ-R.

Overall, LLI teachers were most likely to “Agree” or “Strongly Agree” that they understood the goals of LLI, had received adequate professional development for implementing LLI, had a thorough understanding of how to implement LLI, had instructional materials readily available, and were in a school that believes all children can learn to read and write (all 100%). However, LLI teachers were most likely to “Disagree” or “Strongly Disagree” that LLI teachers are given sufficient planning time to implement the program (60%), their school has sufficient faculty and staff to provide LLI to all students who need the intervention (40%), their administration protects time needed for LLI (40%), and their students' parents participate in LLI home literacy activities (40%).

Table 13: LLITQ Response Frequencies (Fall 2015 $n = 5$)

Item	Percent Responded		
	Strongly Agree/Agree	Neutral	Disagree/Strongly Disagree
	Fall 2015	Fall 2015	Fall 2015
I understand the goals of the Leveled Literacy Intervention (LLI) program.	100.0	0.0	0.0
I have received adequate professional development for implementing LLI.	100.0	0.0	0.0
I have a thorough understanding of how to implement LLI.	100.0	0.0	0.0
Guidance and support are provided by our instructional and administrative staff to help us implement LLI.	80.0	20.0	0.0
I believe LLI has positively impacted LLI students' literacy achievement.	80.0	20.0	0.0
LLI teachers are given sufficient planning time to implement the program.	40.0	0.0	60.0
Students who receive LLI in this school are more enthusiastic about reading, writing, and learning because of LLI.	60.0	40.0	0.0

Table 13: Continued

Item	Percent Responded		
	Strongly Agree/Agree	Neutral	Disagree/Strongly Disagree
	Fall 2015	Fall 2015	Fall 2015
Our school has sufficient faculty and staff to provide LLI to all students who need the intervention.	20.0	40.0	40.0
Our administration protects the time needed for daily uninterrupted LLI teaching.	20.0	40.0	40.0
Our students' parents participate in LLI home literacy activities with their child(ren).	20.0	40.0	40.0
Teachers in this school are generally supportive of LLI.	80.0	0.0	20.0
Ongoing communication exists between LLI teachers and classroom teachers.	40.0	40.0	20.0
LLI teachers are encouraged to communicate concerns, questions, and constructive ideas regarding the program to school staff or administration.	60.0	40.0	0.0
LLI allows for teachers to provide differentiated instruction to address the varying strengths and needs of students.	80.0	20.0	0.0
Instructional materials (books, assessments, and other resources) needed to implement LLI are readily available.	100.0	0.0	5.0
The faculty, staff, and administration in my school believe that all children can learn to read and write.	100.0	0.0	0.0
LLI is aligned with state and district reading and language arts standards.	80.0	20.0	0.0
LLI training has improved my reading instruction.	80.0	20.0	0.0
LLI students perform better on state assessments as a result of their participation in LLI.	40.0	40.0	20.0
Because of LLI, I have a greater understanding of...			
The reading process.	80.0	20.0	0.0
The characteristics of leveled books and their relationship to successful reading.	80.0	20.0	0.0
The role of comprehension in successful reading.	80.0	20.0	0.0
How to improve children's writing strategies.	60.0	40.0	0.0
Item	Percent Respondent		
	Extensively/Sufficiently	Somewhat	Not At All
	Fall 2015	Fall 2015	Fall 2015
To what degree does your school administration support your efforts as an LLI teacher?	40.0	60.0	0.0
To what degree does the district support your efforts as an LLI teacher?	80.0	20.0	0.0
To what degree does your teaching schedule allow time to implement LLI effectively?	40.0	60.0	0.0
To what extent do you feel LLI has helped your English Language Learner students?	20.0	20.0	20.0
To what extent do you feel LLI has helped your students with special needs?	40.0	40.0	20.0

Table 13: Continued

Item	Percent Respondent		
	Regularly (Every day)/ Frequently (3-4 days per week)	Occasionally (1-2 days per week)	Rarely (Less than 1 day per week)/Not At All (Never)
	Fall 2015	Fall 2015	Fall 2015
How often did your LLI group lessons last 45 minutes or more?	80.0	0.0	20.0
Were you able to meet every day with your LLI group(s)?	80.0	20.0	0.0
How often did you follow the LLI lessons exactly as instructed in the Lesson Guide?	80.0	20.0	0.0
How often were you able to implement LLI reading activities?	80.0	20.0	0.0
How often were you able to implement LLI writing activities?	40.0	60.0	0.0
Item	Percent Respondent		
	Fall 2015		
Do you think your school should continue the Leveled Literacy Intervention program?			
Yes	100.0		
No	0.0		

Note. Item percentages may not total 100% due to missing input from some participants.

The LLITQ invited LLI teachers to share open-ended comments regarding the reasons their schools should continue or not continue using the LLI system, and the strengths and areas for improvement of LLI. Participants’ responses to these items were summarized using a structured, multi-step process. First, the original comments were assigned codes representing their basic content. Next, these codes were grouped into categories, which were then organized into overarching themes. Final analysis produced frequency percentages for each theme. Because it was possible for some comments to contain multiple content codes, the percentages reported reflect the total number of codes within each theme and not necessarily the total number of comments received from participants. Responses for each question are summarized and discussed below.

Continuation of LLI

In Fall 2015, when participating LLI teachers in SPS district were asked why their school should continue or not continue using LLI, no teachers provided a reason for their schools to discontinue the use of LLI. Responses regarding the reasons to continue mainly focused on the positive effects for students and teachers, particularly student engagement. One teacher stated that her students really enjoyed the LLI books.

In Spring 2016, when participating LLI teachers in SPS district were asked why their school should continue or not continue using LLI, the majority of teachers (83.3%) provides reasons why their school should continue LLI. Reasons to continue LLI most frequently focused on LLI’s effectiveness in improving literacy skills (66.7%). Another reason cited by LLI teachers involved the program design and instruction (16.7%), particularly LLI’s intervention design for low achieving students. Only 16.7% teachers’ comments indicated that they might need more time to implement LLI properly.

Strengths of LLI

In Fall 2015, when participating LLI teachers in SPS were asked about the strengths of LLI, one third of their responses focused on the materials and resources – particularly the colorful and interesting books. An additional one third of teachers’ responses involved the instructional components, particularly word work. Finally, in one third of overall responses regarding the LLI strengths, the participating teachers discussed the design and organization of the LLI system, particularly the regular basis of LLI.

In Spring 2016, when participating LLI teachers in SPS were asked about the strengths of LLI, half of their overall responses focused on the materials and resources – particularly the interesting books, which were mentioned in all the responses. The next most common theme regarding strengths involved the instructional components of LLI (33.3%); these responses most frequently cited the word work (50.0%), the assessment (25.0%) and the feedback components (25.0%) of the LLI system. Lastly, in 16.7% of overall responses, the participating teachers discussed the design and organization of LLI system. Of these comments, teachers referenced the fact that the lesson plans of LLI are comprehensive and easy to follow.

Improvements for LLI

In Fall 2015, regarding areas for improvement of the LLI system, all participating LLI teachers in SPS discussed the program implementation. With this theme, all teachers stressed the need for more time during lessons. Teachers stated that there is a lot to get done in a lesson, so they need more time to fully utilize the 45 minutes.

In Spring 2016, regarding areas for improvement of the LLI system, participating LLI teachers in SPS most frequently discussed the program implementation (60.0% of overall comments). With this theme, all teachers highlighted the need for more time during lessons. Another theme that frequently arose regarding areas for improvements involved specific strategies and instructional components (20.0%), particularly the homework. Teachers mentioned that more coordination is needed between parents, LLI teachers, and classroom teachers to ensure the homework is completed more consistently. Lastly, 20.0% of teachers’ overall comments were related to LLI materials and resources, specifically the books (e.g., a greater scope of experience should be covered in the books).

Sample comments from LLI teachers are provided below.

Reasons to Continue or Not Continue

“We should continue LLI because it truly shows progress in students who are struggling or need intervention to get closer to or on grade level.”

Strengths of LLI

“Students are provided with many opportunities to read, discuss, and write about fiction and nonfiction texts with their teacher and peers. They participate in word study and complete frequent assessments. They receive constructive feedback that helps them to improve and succeed.”

Areas for Improvements

“Realistic times for the writing activities. Many books seem to be from New England so maybe a greater scope of experiences other than by the Sea, or Sea life. Many kids wanted football and hockey.”

Classroom Teacher Literacy Instruction Questionnaire (CTLIQ)

The Classroom Teacher Literacy Instruction Questionnaire (CTLIQ) was administered online to regular classroom teachers during Fall 2015 and Spring 2016 as a general measure of classroom teachers’ literacy instructional strategies and perceptions of LLI and the core literacy program at their schools. However, because of a particularly small sample size for the Fall, results are only reported for the Spring in order to best maintain confidentiality (Spring $n=7$). Table 13 illustrates the frequencies of responses for each item on the CTLIQ.

Overall, participating classroom teachers were most likely to “Agree” or “Strongly Agree” that they have received adequate professional development for implementing the school’s core literacy program, they have a thorough understanding of how to implement the school’s core literacy program, they believe the program has positively impacted students’ literacy achievement, the school believes all children can learn to read and write, the core program is aligned with state standards, and professional development for the core program as improved reading instruction (all 100%). Also, classroom teachers “Regularly” or “Frequently” reported that students participate in whole group reading instruction (85.7%), students participate in writing activities (100%), they integrate both vocabulary and comprehension into their literacy instruction and activities (71.4%), they read high-quality literature to their students and engage in discussions about the text (100%), and they assign students home literacy activities (85.7%). Regarding less positive areas of their school’s core literacy program, classroom teachers reported they “Disagree” or “Strongly Disagree” that teachers are given sufficient time to implement the core literacy program (28.6% in Spring). All items can be found in Table 14 below.

Table 14: CTLIQ Response Frequencies (Spring 2016 $n = 7$)

Item	Percent Responded		
	Strongly Agree/Agree	Neutral	Disagree/Strongly Disagree
	Spring 2016	Spring 2016	Spring 2016
I understand the goals of our school's core literacy program.	85.7	14.3	0.0
I have received adequate professional development for implementing our school's core literacy program.	100.0	0.0	0.0
I have a thorough understanding of how to implement our school's core literacy program.	100.0	0.0	0.0
Guidance and support are provided by our instructional and administrative staff to help us implement our core literacy program.	85.7	14.3	0.0
I believe our core literacy program has positively impacted students’ literacy achievement.	100.0	0.0	0.0
Teachers are given sufficient planning time to fully implement our school's core literacy program.	28.6	42.9	28.6
Students in this school are more enthusiastic about reading, writing, and learning because of our core literacy program.	71.4	28.6	0.0

Table 14: Continued

Item	Percent Responded		
	Strongly Agree/Agree	Neutral	Disagree/Strongly Disagree
	Spring 2016	Spring 2016	Spring 2016
Our school has sufficient faculty and staff to fully implement its core literacy program.	42.9	42.9	14.3
Our administration protects the time needed for daily uninterrupted core literacy instruction.	71.4	14.3	14.3
Our students' parents participate in home literacy activities with their child(ren).	42.9	28.6	28.6
Teachers in this school are generally supportive of our core literacy program.	85.7	14.3	0.0
Teachers are encouraged to communicate concerns, questions, and constructive ideas regarding our core literacy program to school staff or administration.	71.4	14.3	14.3
Our core literacy program allows for teachers to provide differentiated instruction to address the varying strengths and needs of students.	85.7	14.3	0.0
Instructional materials (books, assessments, and other resources) needed to implement our core literacy program are readily available.	85.7	14.3	0.0
The faculty, staff, and administration in my school believe that all children can learn to read and write.	100.0	0.0	0.0
Our core literacy program is aligned with state and district reading and language arts standards/frameworks.	100.0	0.0	0.0
Professional development for our school's core literacy program has improved my reading instruction.	100.0	0.0	0.0
Our core literacy program adequately prepares our students for state assessments.	100.0	0.0	0.0
I have a clear understanding of the Leveled Literacy Intervention (LLI) program.	42.9	14.3	42.9
LLI supports the goals of my school's core literacy program.	57.1	42.9	0.0
Ongoing communication exists between LLI teachers and classroom teachers.	42.9	42.9	14.3
Students who participate in LLI show increased enjoyment of reading and writing.	57.1	14.3	28.6
Students who participate in LLI show increased achievement in literacy.	57.1	28.6	14.3
Students who participate in LLI show increased participation in classroom literacy activities and instruction.	42.9	42.9	14.3
Item	Percent Responded		
	Extensively/Sufficiently	Somewhat	Not at all
	Spring 2016	Spring 2016	Spring 2016
To what degree does your school administration support your efforts to implement your school's core literacy program?	100.0	0.0	0.0
To what degree does the district support your efforts to implement your school's core literacy program?	100.0	0.0	0.0
To what degree does your teaching schedule allow time to implement your school's core literacy program effectively?	71.4	28.6	0.0
To what extent do you feel your school's core literacy program has helped your English Language Learner students?	28.6	57.1	0.0
To what extent do you feel your schools' core literacy program has helped your students with special needs?	42.9	42.9	14.3

Table 14: Continued

Item	Percent Responded		
	Regularly (Every day)/ Frequently (3-4 days per week)	Occasionally (1-2 days per week)	Rarely (Less than 1 day per week)/Not At All (Never)
	Spring 2016	Spring 2016	Spring 2016
Students participate in whole group reading instruction.	85.7	0.0	14.3
Students participate in small group or individual reading instruction.	57.2	42.9	0.0
I provide guided reading instruction using leveled texts for groups of students with similar learning levels.	57.2	42.9	0.0
Students meet in small, heterogeneous groups to discuss the books that they are reading.	57.2	28.6	14.3
Students participate in writing activities, such as mini-lessons, independent writing, conferencing, and sharing.	100.0	0.0	0.0
I provide opportunities to develop oral reading fluency (e.g., shared reading, partner reading).	57.2	28.6	14.3
I teach phonological awareness (sound patterns, rhymes, etc.) to my students.	28.6	28.6	42.9
I integrate both vocabulary and comprehension into my literacy instruction and activities.	71.4	28.6	0.0
I read high-quality children’s literature (e.g., fiction, non-fiction, poetry) to my students and engage them in interactive discussion.	100.0	0.0	0.0
I assign students home literacy activities to encourage parent participation.	85.7	0.0	14.3
Item	Spring 2016		
Do you think your school should continue the current core literacy program?			
Yes	100.0		
No	0.0		

Note. Item percentages may not total 100% due to missing input from some participants.

The CTLIQ also invited classroom teachers to share open-ended comments regarding the strengths and areas for improvement of their school’s core literacy program and the reasons that their school should continue or not continue the core literacy program. Classroom teachers’ open-ended responses were analyzed using the same structured, multi-step process employed for the LLITQ comments and are summarized by question below. Responses for each question are summarized and discussed regarding each question below.

Continuation of Core Literacy Program

In Spring 2016, participating SPS classroom teachers were asked why their school should continue or not continue using the current core literacy program. 100% of the comments shared by the respondents believed the program should be continued. One fourth of their comments supported the reading component of the program (25.0%) while one fourth focused on the program’s differentiated instruction (25.0%). Other categories within this theme supported the writing component (18.7%), and the belief that it is effective (18.7%), regarding the progress and growth of students.

Strengths of Core Literacy Program

In Spring 2016, participating SPS classroom teachers were asked about the strengths of their school's core literacy components. Of the overall responses, nearly half (47.6%) focused on the resources. A third of the resources mentioned were support for teachers (30.0%), followed by books (20.0%) and professional development (20.0%). Of the responses, 28.5% focused on the instructional components, half mentioned were focused on literacy skills while a third of respondents mentioned the reading portion (33.3%). Additionally, one teacher commented on teacher implementation (16.6%). The responses also found positive program characteristics to be a strength of the core literacy program. These categories consisted of the ability to meet individual students' needs (60.0%), alignment with standards (20.0%), and ample amount of time (20.0%).

Improvements for Core Literacy Program

Regarding areas of improvement for the core literacy program, Spring 2016 SPS teachers most frequently commented on the instructional component with 60.0% of the overall comments. Within this theme, a third (33.3%) of the comments were related to writing skills, a third (33.3%) related to the need for more focus on literacy skills, and a final third (33.3%) focused on time restraints. Further, classroom teachers cited the curriculum organization and delivery to be an area of improvement in 40.0% of the comments, in which all comments focused on organization.

Sample comments from LLI teachers are provided below.

Reasons to Continue or Not Continue the Core Literacy Program

"The current program is growing and developing at our school. I believe that it is beneficial to all learners. It provides an opportunity for slower learners to gain skills at their level, while also providing for opportunities for the gifted student to develop at a faster pace on their level."

Strengths

"The great majority of students love reading and writing. They enjoy learning, discussing their reading and writing, and sharing their opinions. The writing and reading workshop format of mini lesson, student action regarding the mini lesson, and sharing is such a simple but powerful design."

Areas of Improvement

"Programs need to be fully investigated and researched before teachers are expected to implement. We must have ample training and guidance to do what is asked from the district. This training must be done before the school year starts and all materials must be available in enough time for teachers to plan ahead."

Leveled Literacy Intervention Principal Questionnaire (LLIPQ)

Given there was only one responding principal for SPS, the LLIPQ responses are only able to be reported in the aggregate across all three districts to protect confidentiality. Please see the overall report for LLIPQ results from principals across all three districts.

Home Literacy Support Questionnaire (HLSQ)

The Home Literacy Support Questionnaire (HLSQ) was administered to parents/guardians of treatment and control group students at the end of the school year as a general measure of their support for literacy at home and perceptions of their child’s literacy instruction at school – including LLI, if applicable ($n = 48$). Table 15 illustrates the frequencies of responses for each item on the HLSQ. Most of the participating parents/guardians reported positive perceptions of their child’s literacy activities at home and school and the amount of home literacy support they provide. Further, of those parents/guardians who indicated that their child has participated in LLI, most shared positive perceptions of the experience.

Overall, participating parents/guardians were most likely to “Agree” or “Strongly Agree” that they encourage their child to practice reading at home (97.9%), that they have books at home for their child to read (95.8%), and that they believe their child can become a good reader and writer (93.8%). Additionally, of those parents indicating that their child participated in LLI, approximately three-quarters of parents “Agreed” or “Strongly Agreed” that their child’s school should continue using LLI (78.3%) and that their child’s participation in LLI improved his/her reading and writing (78.3%). Over half of the participating parents/guardians “Agreed” or “Strongly Agreed” that they participate in LLI take-home activities with their child (65.2%). All items can be found in Table 15 below.

Table 15: HLSQ Response Frequencies (n = 48)

Item	Percent Responded		
	Strongly Agree/Agree	Neutral	Disagree/Strongly Disagree
My child enjoys reading and writing.	79.2	8.3	12.5
My child reads and writes at home.	83.3	8.3	6.3
I read and write with my child at home.	79.2	16.7	4.2
I have books at home for my child to read.	95.8	0.0	2.1
I read books to my child at home.	68.8	18.8	12.5
I encourage my child to practice reading at home.	97.9	2.1	0.0
I encourage my child to practice writing at home.	79.2	20.8	0.0
I believe my child can become a good reader and writer.	93.8	4.2	0.0
I am pleased with the instruction my child is receiving in reading and writing at school.	89.6	6.3	4.2
I participate in reading and writing activities at my child’s school.	39.6	22.9	37.5
I know how my child is doing in reading and writing at school.	87.5	8.3	4.2
Item	Percent Responded		
Has your child participated in the Leveled Literacy Intervention (LLI) program at his/her school?			
No	25.0		
Yes	47.9		
Not sure	27.1		

Table 15: Continued

Item	Percent Responded		
	Strongly Agree/Agree	Neutral	Disagree/Strongly Disagree
I think my child’s participation in LLI has improved his/her reading and writing.	78.3	21.7	0.0
I participate in LLI take-home activities with my child.	65.2	8.7	26.1
I think my child’s school should continue using the LLI program.	78.3	17.4	4.3

Note. Item percentages may not total 100% due to missing input from some participants.

Parents/guardians who completed the HLSQ also responded to open-ended questions regarding their perceptions of the strengths and areas for improvement of their child’s literacy instruction at school. Their open-ended responses were analyzed using the same structured, multi-step process employed for the teacher and principal surveys and are summarized by question below.

When asked their opinion of the best things about the reading and writing instruction their child receives at school, participating parents/guardians most commonly discussed the instructional components (34.9%), including: books/materials (53.3%), writing activities (26.7%), group work (13.3%), and one-on-one or small group instruction (6.7%). In addition, approximately one-third (32.6%) of responses related to the positive impact the program had on the students, including: children having learned and/or noticeably improved (42.9%), children having more enthusiasm for reading and writing (35.7%), and children having more confidence regarding reading and writing (21.4%). Further, slightly more than one quarter (27.9%) of the overall responses were related to positive characteristics of the program in general, including: instruction and experiences that are enjoyable (41.7%), instructions and expectations that are clear (25.0%), encouragement to read and write (16.7%), support for struggling students (8.3%), and opportunity for home practice (8.3%). Finally, in 2.3% of comments, respondents generally discussed liking the literacy program and feeling pleased with the instruction their child is receiving, though respondents in an equal percent of comments (2.3%) stated that they were not sure or did not share any positive perceptions.

Participating parents/guardians were also asked what changes they would like to see in their child’s reading and writing instruction. The most common response was that no changes were needed or that they were pleased with their child’s progress (34.5%). In slightly less than one-third of responses (31.0%), participants discussed the need for additional support or resources, including: communication with parents/guardians (33.3%), school-level resources and support (e.g. expanded library access, classroom books, expansion of the literacy program across grades, etc.; 33.3%), individualized attention for students (11.1%), emphasis on making reading and writing interesting and enjoyable (11.1%), and additional take-home activities and materials (11.1%). Further, in slightly more than one quarter (27.6%) of responses participants recommended instructional changes, including: greater emphasis on writing and/or handwriting (62.5%), greater emphasis on spelling and grammar (25.0%), and more challenging instruction (12.5%). Finally, 3.5% of respondents stated not being sure what could be improved. Another 3.5% of comments were not directly related to improvements for the reading program.

Sample comments are provided below.

Strengths

"I believe the instruction he has received this year has helped him become a better reader. I see him comprehending more. He enjoys reading much more now."

Changes Needed

"More information for parents regarding the format/instruction for writing. More grammar instruction across all grades is needed."

Stakeholder Feedback

LLI Teacher Focus Group

In order to obtain feedback regarding implementation of the LLI system from current LLI instructors, a voluntary focus group was conducted with those LLI teachers who took part in the study. A semi-structured focus group protocol was utilized. Responses are summarized by question below.

Responses from LLI teachers were overwhelmingly positive in regard to their overall perceptions of LLI. Respondents reported loving LLI because the program was well-developed and research-based. Because of this the teachers stated that the components work in general, and one teacher commented that it is clear there is a reason behind everything one is asked to do. Another teacher mentioned that "how the program is developed almost guarantees progress". Teachers also mentioned that the students love the program because they think it is fun. One teacher commented that the students "love to come and some will stay late by choice". In addition, all of the teachers mentioned seeing growth in their students, not only in scores/grades but also in confidence and love of reading.

When LLI teachers were asked to discuss the strengths of the LLI system, responses were varied and enthusiastic. One of the most frequently mentioned strengths was the material, including the books and activities. Teachers thought that there was a good variety of books that were interesting to the students. They also mentioned that the students liked the activities, such as the word study, because they "feel like it is a game; they didn't know they were doing work". Many of the teachers also thought a strength of the program was that it helped develop a sense of community that made the students feel comfortable and confident. In this regard, some of the teachers mentioned that the students formed a close-knit group that increased their comfort in reading and volunteering to participate in activities. One teacher mentioned that this was because the students were similar in their level of reading skill and that the confidence built within the group sometimes made students more likely to participate in regular classroom activities, like reading out loud. A last strength commonly mentioned was the balance of the program. Teachers expressed general contentedness with how the program tapped into multiple areas, including reading, writing, and speaking.

When asked about areas of LLI that might need improvement, LLI teachers primarily mentioned the amount of content that needs to be covered, the complexity of some writing assignments, and the dependence on homework. In terms of content, teachers mentioned that there was too much content that needed to be covered for some lessons and that this could not fit into a 45-minute session. Related to this, some of the teachers also believed that some of the writing assignments were too complicated and that this also was difficult to fit into a short length of time. One teacher mentioned that writing

assignments often had to be adjusted so that it was still covered while still being able to get to other parts of the lesson. Finally, one of the teachers mentioned that the dependency on homework was an issue. This teacher stated that the program assumes that students will be doing the homework but that this is difficult to ensure, particularly when the students already have other homework from their regular classes.

When focus group participants were asked to comment on common logistical issues they encountered when implementing the LLI system, many comments focused on issues with scheduling and consistency in getting students, the amount of time allotted, and difficulties having students of significantly different reading level in the same group. In terms of scheduling, respondents commented that it was sometimes hard to secure a space for the groups and that getting the students for the groups on a regular basis was difficult. Consistency in getting the students for group was seen as the largest of these concerns, as students were sometimes late, thus not receiving the full 45 minutes, or would be held in the regular classroom to complete other assignments that were considered more important by the regular classroom teachers. One respondent mentioned that some classroom teachers have even sent assignments for the students to finish before returning from the LLI group, preventing actual participation even when the student is there. A final logistical issue mentioned was the difficulty of effectively addressing great differences in student reading levels within the same group. One teacher mentioned that having quicker readers in with slower readers prevented the groups from being most effective. In this case, the teacher mentioned that it would be helpful to have the students with others of similar level. However, the teachers mentioned that rotating students to other groups was not always feasible given the limited number of groups and teachers.

When LLI teachers were asked how effective LLI has been in meeting individual students' needs, they generally reported that it was effective, but that there were some caveats. Teachers reported seeing progress in many of their students, however they stated that even more progress may have been possible if they could have re-grouped both the students that were struggling and that were making significant progress. Teachers mentioned that not being able to re-group students by similar reading levels influenced the pace which impacted student progress. Teachers mentioned that re-grouping had been happening in the previous year but that it did not happen this year. Some teachers also mentioned that they wish they had been able to release some students from the group in order to bring in others that needed more help. Related to these issues of student grouping, many of the teachers mentioned issues with placing students with special needs in with the regular LLI groups. One issue mentioned was that these students were more likely to feel overwhelmed in the groups and subsequently feel uncomfortable or lose interest. In addition to this issue, there were some issues about the content of the education plans for students in special education, as many of these plans did not have reading goals. One teacher also mentioned that LLI might not be appropriate for students with decoding problems or other such issues. There was also an additional issue with scheduling for students in special education, as they followed a schedule that was often different from the other students, since they had a number of other services they needed to receive as well.

LLI teachers reported mixed support from administration regarding LLI. While some teachers felt very supported by LLI, others commented that administration had taken their rooms away or called on them to substitute without notice. The teachers mentioned that most of these were issues of communication, but that the administration also had mixed perceptions of the value of LLI, mostly due to differences in how much they knew about the program. There were also mixed levels of support from and communication with the literacy coaches at the schools. When asked about other school faculty, participants reported mixed support from classroom teachers. Respondents reported that, while a

number of teachers were supportive, others did not know much about LLI and felt that the students were missing something from their class. Because of this, some had issues getting the teachers to consistently send the students to the LLI groups. Many of these issues were also contributed to a lack of communication and availability of the faculty to talk and create schedules.

When asked their opinions about the training they received, many of the LLI teachers reported that the training was generally very effective and supportive, though there were some areas for improvement. Respondents mentioned that most of the people attending the training saw a benefit in it and were able to get helpful information. One respondent mentioned an intent to use the information from LLI from now on, commenting on having had “the most productive two years” after learning about LLI. Respondents also mentioned that the training is particularly helpful for new teachers using LLI. In terms of improvement, respondents mentioned wanting more time to openly discuss questions with others, particularly for those that already have background knowledge of the program. In this regard, many of the respondents thought that it was more helpful to discuss suggestions and potential solutions to specific issues than to have set topics and lecture-style training each day.

Site Researcher Focus Group

Because the site researchers who collected observational and student benchmark data for the study were primarily retired teachers who had experience teaching in the districts, CREP researchers utilized focus groups to solicit their feedback regarding LLI and its implementation. The site researchers were able to provide an objective “outsider’s” perspective based on their random observations of the LLI groups. Responses are summarized below.

When asked their overall perceptions of LLI, the on-site researchers’ responses were very positive. One respondent commented that, “...it is an amazing program, it is worth every minute they spend in the program”. Another respondent also agreed with this, stating that LLI has a great deal of potential for student learning.

On-site researchers also shared their perceptions of the strengths of LLI. The biggest strength mentioned by respondents was the group structure and individual focus of the program. In this regard, respondents mentioned that having small groups was “crucial and powerful”. Respondents also reported that being able to group students by ability was very beneficial. One respondent commented that it was “like having a guided reading group every day”. It was also mentioned that the structure of the grouping allowed LLI teachers to push students further when needed.

When asked what areas of LLI may need improvement, respondents mentioned issues with having time to cover all content and follow the lesson plans. It was mentioned that the teachers were not able to cover all of the content required in each lesson, so they had to pick and choose what to cover. Because of this issue, the lesson plans were not always being followed.

Participants in the focus group were also asked about student responses during LLI observations and the quality of LLI for the study students. In regard to student responses to LLI, respondents commented that the students were always actively engaged and that the students liked being in LLI. Respondents saw this engagement in and enjoyment of LLI as an indicator of quality in both the teachers and the program. In regard to the quality of LLI for the students, respondents mentioned that the quality was overall good, but that it could be increased further by additional training. Respondents also

mentioned being bothered by the lack of writing during LLI. Respondents frequently saw word lessons and games, but they did not see much if any writing occurring.

Finally, participants were asked about their perceptions of the training they received and the instruments they used to collect the data for the study (i.e., the LOT and LLIOT observations and the Fountas & Pinnell Benchmark Assessment System). Respondents reported that the training was overall very good with a few areas for improvement. One respondent mentioned feeling uncomfortable with the observations, stating that more training on what LLI should look like would be helpful in making observations better. However, another respondent mentioned that this may relate to the amount of experience one already has with LLI, as they felt comfortable with all aspects of the training. In terms of the instruments uses, the respondents reported that the instruments were very comprehensive and that they did not have any criticisms.

Conclusions

1. What progress in literacy achievement, if any, do students who receive LLI in grades 3-5 make compared to students who receive core literacy instruction alone?

Results revealed that LLI had a positive impact on some 3rd-5th grade students' literacy achievement. In particular, one of the three types of analyses showed important results: When the control group had an advantage at baseline and the treatment still had an advantage in outcomes. While none of these positive effects were statistically significant, some were considered substantively important based on guidelines from the What Works Clearinghouse (WWC) (i.e., an effect size of +/- 0.25; What Works Clearinghouse, 2014). Detailed findings for each subgroup are presented in the body of this report; however, some of the strongest findings are outlined below.

With regard to **benchmark levels**:

- When starting at a disadvantage, 3rd grade students overall showed substantively higher gains compared to control students.
- When starting at a disadvantage, 4th grade students overall, as well as high achieving students, showed substantively higher gains compared to control students.
- In addition, when starting at a disadvantage, 3rd grade high achieving students and 5th grade students overall showed higher gains compared to control, but these did not qualify as substantively important.

2. At what level of fidelity to the program model is LLI implemented by teachers participating in the study?

Overall, the observation results from the current study suggest that LLI was implemented with inconsistent fidelity to design. When observed, many lesson components received acceptable to high fidelity ratings, with few indications of needing improvement. However, a number of components went unobserved in over half of the observations. In addition, a few areas were rated as needing improvement, namely introducing vocabulary words (30.4%), having well organized lessons (34.8%), and appropriately pacing lesson components (56.5%). Further, the majority of site researchers concluded that, overall, the lessons needed improvement, as they were not delivered as designed (60.9%). However, site researchers had a more neutral opinion in their open-ended comments, of which very few indicated specific areas needing improvement. Additionally, the observation results revealed that LLI

implementation was mostly consistent over the school year at both time points when the observations were conducted. Changes in implementation over the year were only noted for 5th grade, where there was substantively important improvement in quality of LLI instruction for 5th grade but slight decreases in the areas of literacy instructional strategies and learning environment. In general, it should also be noted that, for 3rd grade, all subscales were rated between needing improvement and acceptable at both time points.

The observation results were complemented with self-report feedback from the participating LLI teachers, which showed a slightly different picture. In particular, large majority reported implementing LLI as designed (e.g., meeting daily for 45 minutes, following the LLI Lesson Guide), understanding the LLI goals and procedures, and having sufficient training to implement LLI effectively. However, LLI teachers did note that they were often asked to do other tasks that conflicted with LLI lesson time, feeling that they did not have sufficient planning time to implement the program and that administration did not always protect the time needed for uninterrupted teaching. This may have impacted teachers' implementation and students' progress as seen in the overall achievement results.

Finally, the LLI attendance records that were available (97% of treatment group) from the current study revealed that there were, on average, 67 LLI instructional days, compared to the recommended number of LLI instructional days (i.e., approximately 90-120 days/18-24 weeks). Of these students with attendance data, when looking individually at each student, the data revealed that none of the treatment students received the recommended dosage. Student absences were due to several student-level factors (e.g., individual absences or unavailability during LLI group time) as well as school or district limitations (e.g., holidays, assessment windows during which LLI teachers and/or students were pulled during LLI group time, delays in starting LLI due to scheduling conflicts or difficulty accessing student data). Therefore, the findings at each grade level which are not meeting statistical significance or substantively important progress may have been impacted due to a large number of treatment students not receiving a full dosage of LLI. Schools should note the importance of consistently providing LLI throughout the year so the students can make the most progress by receiving, at a minimum, the recommended amount of LLI lessons.

3. What are stakeholders' perceptions of the LLI system for grades 3-5 and the core literacy program?

Overall, LLI teachers, classroom teachers, principals, parents/guardians, and site researchers shared positive perceptions of the LLI system and its impact on struggling students' literacy success. Stakeholders felt that LLI has benefits for students' literacy achievement and skills as well as their engagement, interest, and confidence related to reading and writing. Stakeholders also reported positive perceptions of such aspects of the LLI system as its design, instructional components, and materials (particularly the lesson books). However, although stakeholders generally perceived LLI as helpful, there was common feedback regarding the need for improving the logistics of implementing LLI. In general, stakeholders discussed the need for more time to complete lessons and for that time to be protected for LLI teaching, as well as needs for better scheduling and more staff in order to appropriately serve students.

Regarding the core literacy instruction, stakeholders' perceptions were generally positive, although some areas of concern were identified. Stakeholders perceived that their schools are generally supportive of literacy and provide a high-quality learning environment conducive to learning. Further, stakeholders shared positive perceptions of the core literacy program's impact on student engagement,

as well as such aspects of the program as classroom materials. The presence of guided reading, independent literacy work, and small, as well as whole group instruction were also noted as being positive aspects of the core literacy programs. However, stakeholders agreed that the core literacy instruction also has areas of improvement. Areas for improvement suggested by stakeholders included more time for literacy, better scheduling, clearer communication of details about the core literacy program, and an increased emphasis on technical reading and writing skills, including grammar.

Limitations of the Study and Recommendations

Although the current study produced some important positive findings regarding the efficacy of LLI, particularly in 3rd and 4th grade, several factors were encountered that may limit the generalizability of the findings and that prevented researchers from obtaining adequate power to draw definitive conclusions in some areas. These limitations are summarized below, followed by data-based recommendations for improvement.

The primary limitation facing the current study involved the sample being non-randomized, which, while not ideal for research, was a real-world constraint for obtaining districts that would participate in the study. While the study had a control group of students who were matched on demographic characteristics and initial reading levels, we also conducted baseline equivalence testing to determine if the treatment and control groups were equivalent groups at the beginning of the study. Our results found, in several cases, that they were not equivalent at the start, and because of this, our findings in those instances (as noted in the results section) should be interpreted with caution.

The sample size, or the number of students that were able to participate in the study, of our subpopulations was also a limitation. Despite efforts to select districts with diverse populations, the study still was limited in the numbers of participants in ELL, Special Education, and some ethnicity groups. Although some positive effects were detected with marginal statistical significance and/or “substantively important” effect sizes according to What Works Clearinghouse guidelines (U.S. Department of Education, 2011), the small size of some subgroup samples may have made it more difficult to detect significant differences between the literacy gains of treatment and control group students.

A third limitation of the study design was the fact that control group students were allowed to receive other supplemental literacy services while they were participating in the study, as long as they did not receive LLI until after the study was over. This was a district-level request that was necessary in order for them to agree to participate in the research study. Additionally, the supplemental literacy services received by control group students could vary from school to school and district to district. The fact that both groups could receive additional literacy services may have resulted in a smaller difference in literacy gains between treatment and control group students. However, this limitation serves to increase the meaningfulness of the significant gains made by treatment group students in comparison to the control group, because receiving LLI helped these treatment group students outperform control group students who not only received core literacy instruction, but also supplemental literacy services.

A final limitation of the study was the fact that treatment group students did not receive the recommended amount of LLI instructional time as a result of individual absences, delays in starting LLI due to the time required to obtain consent and pretest students, and district-level factors (e.g., holidays, assessment days, and LLI teachers being pulled from their LLI groups for other activities). Although it is not clear whether receiving the recommended number of instructional days would have produced more

significant results, it is possible that maximizing LLI instructional time would have resulted in greater student gains. However, similar to the above limitation regarding control group students receiving supplemental literacy services, this limitation serves to increase the meaningfulness of treatment group students' literacy gains during the shortened timeframe in which they received LLI.

Recommendations

The limitations above reflect the inherent complications in performing educational research across different districts. However, despite these limitations, the current study found educationally meaningful, positive effects of LLI on students' literacy achievement when implemented with sufficient fidelity to the LLI model. Further, stakeholders in these districts – including teachers, administrators, and parents/guardians – were supportive of LLI and perceived positive benefits of the LLI system for their students. Altogether, the results from this research allow us to conclude that LLI has a positive impact on some 3rd-5th grade students' literacy achievement. These results also suggest that paying close attention to program fidelity and continuing implementation of LLI with increased attention to delivering the program as designed would be beneficial and offer an opportunity for research-based recommendations that may enhance the system, future LLI research, and ultimately student achievement. From this research, CREP proposes the following recommendations with regard to LLI and its implementation in schools:

Design

1. Some teachers also mentioned specific aspects of the materials, such as the ease of use of the prompting guide. One improvement mentioned regarded the individualization of material for students. For example, if a student needed more work with phonics or mastered some levels but were not ready for the next highest level, they might have to repeat work, which was not enjoyable for the students.

Implementation

1. In both the current study and a previous study of LLI (Ransford-Kaldon et al., 2010), schools experienced difficulty achieving the minimum of 18 weeks, or 90 days, of LLI instruction. Even though it is possible to implement the program across 18-24 calendar weeks, they are not full weeks of instruction due to holidays, assessments, etc., as well as individual student and teacher absences or unavailability. This is even more difficult during the second semester when schools have multiple end-of-year assessments and activities as well as such events as spring break. Therefore, Districts should prioritize LLI teachers for LLI instruction rather than pulling them for other activities to maximize instructional time for vulnerable students.
2. Districts should be discouraged from allowing students to be pulled for LLI (or other supplemental interventions) during the classroom literacy block.
3. Feedback from LLI teachers suggested that the LLI lessons may be too fast-paced for slower learners, resulting in a lack of time to spend on specific components students need. The district could consider providing recommendations on how best to individualize instruction to meet the

needs of both higher-achieving and lower-achieving students in a group when regrouping is not feasible.

Professional Development (PD) Considerations

1. Some participants felt somewhat overwhelmed by the amount of content and difficulty of the training. Recommendations for improvement focused on providing more time to practice and discuss the material, which participants found to be the most helpful part in learning the nuances of LLI and potential solutions to issues during implementation.
2. Feedback from LLI teachers described examples of high-quality LLI instruction as particularly useful during the LLI PD. The district might consider including more video clips of teachers performing specific LLI instructional routines or strategies during the PD sessions. Additionally, a bank of video clips or webinars on the district website may be helpful in providing ongoing support to LLI teachers.

Future Directions

1. Additional studies empirically evaluating LLI instruction with varying group sizes and varying teaching staff (e.g., Literacy Specialists, Special Education instructors, LLI-trained paraprofessionals) could provide a research-based conclusion as to whether LLI can be adapted to address the limited staff and capacity of some districts preferring to implement LLI with larger group sizes and varying teaching staff.

References

- Alexander, K. L., Entwisle, D. R., & Horsey, C. S. (1997). From first grade forward: Early foundations of high school dropout. *Sociology of Education*, 70(2), 87-107.
- Armbruster, B. B., Lehr, F., & Osborn, J. (2001). *Put reading first: The research building blocks for teaching children to read: Kindergarten through grade 1*. Jessup, MD: National Institute for Literacy.
- Beaver, J., & Carter, M. (2009). *Developmental Reading Assessment, Second Edition: K-8 technical manual*. Upper Saddle River, NJ: Pearson.
- Borman, G. D., Park, S. J., & Min, S. (2015). *The District-Wide Effectiveness of the Achieve3000 Program: A Quasi-Experimental Study*. Online Submission. Retrieved from: <http://files.eric.ed.gov/fulltext/ED558845.pdf>
- Cheung, A. C., & Slavin, R. E. (2013). Effects of Educational Technology Applications on Reading Outcomes for Struggling Readers: A Best Evidence Synthesis. *Reading Research Quarterly*, 48, 277-299.
- Clay, M. (1991). *Becoming literate: The construction of inner control*. Portsmouth, NH: Heinemann.
- Epstein, J., Leung, T., Lloyd, C., Maier, M., & Voorhis, F., (2013). The impact of family involvement on the education of children ages 3 to 8. MDRC. Retrieved from <http://files.eric.ed.gov/fulltext/ED545474.pdf>
- Fountas, I. C., & Pinnell, G. S. (1996). *Guided reading: Good first teaching for all children*. Portsmouth, NH: Heinemann.
- Fountas, I. C., & Pinnell, G. S. (2006). *Teaching for comprehending and fluency: Thinking, talking, and writing about reading, K-8*. Portsmouth, NH: Heinemann.
- Fountas, I. C., & Pinnell, G. S. (2007). *The continuum of literacy learning, grades K-2: A guide to teaching*. Portsmouth, NH: Heinemann.
- Fountas, I. C., & Pinnell, G. S. (2008). *Leveled Literacy Intervention*. Portsmouth, NH: Heinemann.
- Fountas, I. C., & Pinnell, G. S. (2010). *Benchmark Assessment System 1, grades K-2, levels A-N (2nd ed.)*. Portsmouth, NH: Heinemann.
- Harrison, L., Peterman, R., Grehan, A., Ross, S., Dexter, E., & Inan, F. (2008, March). Evaluation of the Leveled Literacy Intervention: Year 1. Paper presented at the annual meeting of the American Educational Research Association, New York.
- Heinemann. (2007). Field study of reliability and validity of the Fountas & Pinnell Benchmark Assessment Systems 1 and 2. Retrieved September 21, 2012, from <http://www.heinemann.com/fountasandpinnell/research/BASFieldStudyFullReport.pdf>

- Heinemann. (2008). Research base for Leveled Literacy Intervention. Retrieved August 29, 2008, from <http://www.heinemann.com/fountasandpinnell/research/LLIResearchBase.pdf>
- Hiebert, E. H., & Taylor, B. M. (Eds.). (1994). Getting reading right from the start. Boston: Allyn & Bacon.
- Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology*, 80, 437-447.
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33, 159-174.
- Lipsey, M. W., & Wilson, D. B. (2001). Practical meta-analysis: Applied social research methods series #49. Thousand Oaks, CA: Sage.
- Lublimer, S. (2004). Help for struggling upper-grade elementary readers. *The Reading Teacher*, 57, 430-438.
- Marcon, R. A. (1995). Early learning and early identification follow-up study: Transition from the early to the later childhood grades, 1990–1993. Washington, DC: District of Columbia Public Schools, Center for System Educational Change.
- National Institute of Child Health and Human Development. (2000a). Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction. Washington, DC: National Institutes of Health.
- National Institute of Child Health and Human Development. (2000b). Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups. Washington, DC: National Institutes of Health.
- Peterman, R., Grehan, A., Ross, S., Gallagher, B., & Dexter, E. (2009, April). An evaluation of the Leveled Literacy Intervention program: A small-group intervention for students in K-2. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.
- Ransford-Kaldon, C. R., Flynt, E. S., Ross, C. L., Franceschini, L. A., Zoblotsky, T. A., Huang, Y., & Gallagher, B. (2010). Implementation of effective intervention: An empirical study to evaluate the efficacy of Fountas & Pinnell's Leveled Literacy Intervention system (LLI). Memphis, TN: University of Memphis, Center for Research in Educational Policy.
- Rashotte C. A., MacPhee, K., & Torgesen, J. K. (2001). The effectiveness of a group reading instruction program with poor reader in multiple grades. *Learning Disability Quarterly*, 24, 119-134.
- Renaissance Learning. (2009). STAR Early Literacy: Technical manual. Wisconsin Rapids, WI: Author.
- Salinger, T. (2003). Helping older, struggling readers. *Preventing School Failure: Alternative Education for Children and Youth*, 47, 79-85.

- Sim, J., & Wright, C. C. (2005). The kappa statistic in reliability studies: Use, interpretation, and sample size requirements. *Physical Therapy*, 85(3), 257-268.
- Sterbinsky, A., & Ross, S. M. (2003). Literacy Observation Tool reliability study. Memphis, TN: University of Memphis, Center for Research in Educational Policy.
- Tabors, P. O., Snow, C. E., & Dickinson, D. K. (2001). Homes and schools together: Supporting language and literacy development. In D. K. Dickinson & P. O. Tabors (Eds.), *Beginning literacy with language: Young children learning at home and school* (pp. 313-334). Baltimore: Brookes.
- Tankersley, K. (2005). *Literacy strategies for grades 4-12: Reinforcing the threads of reading*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Torgesen, J., Stancavage, F., Myers, D., Schirm, A., Stuart, E., Vartivarian, S., Stancavage, F., Durno, D., Javorsky, R., & Hann, C. (2006). *Closing the reading gap: First year findings from a randomized trial of four reading interventions for striving readers*. Washington, DC: Corporation for the Advancement of Policy Evaluation.
- U.S. Department of Education. (2011). *What works clearinghouse procedures and standards handbook* (version 2.1). Retrieved from <http://ies.ed.gov/ncee/wwc/documentsum.aspx?sid=19>
- U. S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. *The Nation's Report Card: Reading Highlights 2003*. Washington, DC: NCES. Retrieved from <http://nces.ed.gov/nationsreportcard/subject/publications/main2013/pdf/2014451.pdf>
- Wanzek, J., & Vaughn, S. (2007). Research-based implications from extensive early reading interventions. *School Psychology Review*, 36(4), 541-561.
- What Works Clearinghouse (2014). *Procedures and standards handbook* (Version 3.0). Washington, DC: Author. Retrieved from ies.ed.gov/ncee/wwc/pdf/reference_resources/wwc_procedures_v3_0_standards_handbook.pdf

Appendices

Appendix A:
Sandwich Benchmark Level Gain Effect Size Summary

Appendix B:
Fountas & Pinnell Grade-Level Equivalence Chart

**Appendix A:
Sandwich Benchmark Level Gain Effect Size Summary**

Subgroup	Grade 3			Grade 4			Grade 5		
	Combined	At/Below Median	Above Median	Combined	At/Below Median	Above Median	Combined	At/Below Median	Above Median
All Students	0.44 [^]		0.21	0.65 [^]		0.64 [^]	0.16		
Economically Disadvantaged Students									
Limited English Proficient Students									
Special Education Students									
African-American Students									
Hispanic Students									
Minority Students									
White Students	0.43 [^]		0.21	0.64 [^]		0.64 [^]	0.16		

Note. **Green** cells are comparisons with baseline equivalence where LLI students had a substantively important or statistically significant advantage on the outcome. **Purple** cells are comparisons where control students had a substantively important advantage on the pretest, but LLI students had an advantage on the outcome. **Pink** cells are comparisons where control students had a substantively important advantage on the pretest, and also had an advantage on the outcome, but the outcome was not substantively important. Cells with an asterisk (*) were statistically significant. Cells with a (^) were substantively important (i.e., effect size (g) ≥ 0.25).

**Appendix B:
Fountas & Pinnell Grade-Level Equivalence Chart**

Grade-Level Equivalence Chart					
Grade	Fountas & Pinnell Level	Basal Level	Reading Recovery Level	Rigby Level	DRA Level
Kindergarten	A	Readiness	1	1-2	A, 1, 2
Kindergarten	B		2	3-4	
Kindergarten					
Grade 1	C	PP1	3, 4	5	3
Grade 1	D	PP2	5, 6	6	4
Grade 1	E	PP3	7, 8	7	6, 7, 8
Grade 1	F	Primer	9, 10	8	10
Grade 1	G		11, 12	9	12
Grade 1	H	Grade 1	13, 14	10	14
Grade 1					
Grade 2	I		15, 16	11	16
Grade 2	J	Grade 2	17, 18	12	18, 20
Grade 2	K		19, 20	13-14	
Grade 2	L			15	24-28
Grade 2					
Grade 3	M			16-17	
Grade 3	N	Grade 3		18	30
Grade 3	O			19	
Grade 3					
Grade 4	P			20	34-38
Grade 4	Q	Grade 4			40
Grade 4	R				
Grade 4					
Grade 5	S				44
Grade 5	T	Grade 5			
Grade 5	U				
Grade 5					
Grade 6	V				
Grade 6	W	Grade 6			
Grade 6	X				
Grade 6	Y				
Grade 7, 8 and Above	Z	Grade 7, 8			

From the 2009 LLI Program Guide