Center for Research in Educational Policy

The University of Memphis
325 Browning Hall
Memphis, Tennessee 38152
Toll Free: 1-866-670-6147

Implementation of Effective Intervention:

An Empirical Study to Evaluate the Efficacy of Fountas & Pinnell's Leveled Literacy Intervention System (LLI)

2009-2010

Center for Research in Educational Policy

The University of Memphis
325 Browning Hall
Memphis, Tennessee 38152
Toll Free: 1-866-670-6147

Implementation of Effective Intervention:

An Empirical Study to Evaluate the Efficacy of Fountas & Pinnell's Leveled Literacy Intervention System (LLI)

2009-2010

September 2010

Carolyn R. Ransford-Kaldon, Ph.D.
E. Sutton Flynt, Ed.D.
Cristin L. Ross, M.S.
Louis Franceschini, Ph.D.
Todd Zoblotsky, Ed.D.
Ying Huang, M.S.
Brenda Gallagher, Ed.D.
Center for Research in Educational Policy

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 evaluate one of their widely used products, the Leveled Literacy Intervention system (LLI), and thank Tift County Schools and the Enlarged City School District of Middletown for working together with us.

Table of Contents

List of Tables

Executive Summary

This report summarizes evaluation results for an efficacy study of the Leveled Literacy Intervention system (LLI) implemented in Tift County Schools (TCS) in Georgia and the Enlarged City School District of Middletown (ECSDM) in New York during the 2009-2010 school year. Developed by Fountas & Pinnell (2009) and published by Heinemann, LLI is a short-term, small-group, supplemental literacy intervention system designed for students in kindergarten through second grade (K-2) who struggle with literacy. The goal of LLI is to provide intensive support to help these early learners quickly achieve grade-level competency.

Both school districts evaluated in this study adopted the targeted, small-group implementation model of LLI in their schools with support from Heinemann consultants providing LLI professional development. This report focuses on the implementation and impact of this model during the first full school year of the system in these schools.

Research Questions

The purpose of this study was threefold: (1) to determine the efficacy of the Leveled Literacy Intervention system (LLI) in increasing reading achievement for K-2 students; (2) to examine the implementation fidelity of LLI; and (3) to determine perceptions of the LLI system according to relevant stakeholders. This study focused on two U.S. school districts and comprised 427 K-2 students who were matched demographically and randomly assigned to treatment and control groups. This evaluation used a mixed-methods design to address the following key research questions:

1. What progress in literacy do students who receive LLI make compared to students who receive only regular classroom literacy instruction?

2. Was LLI implemented with fidelity to the developers' model?

3. What were LLI teachers' perceptions of LLI and its impact on their students' literacy?

Participants

Five elementary schools in TCS in Tifton, GA, and four elementary schools in ECSDM in Middletown, NY, volunteered to participate in the study.[1] TCS is a rural school district located approximately 181 miles south of Atlanta, GA, that served 7,551 students during the 2008-2009 school year. Most of the schools in TCS are small and serve primarily White and African American populations (48.0% and 35.0%, respectively), with more than half of students (65.0%) identified as "economically disadvantaged" by the Georgia Department of Education. Twenty-one K-2 teachers trained in LLI and 209 K-2 students eligible for LLI in TCS participated in this study.

ECSDM is a suburban school district located approximately 72 miles northwest of New York City, NY, that served 6,764 students during the 2008-2009 school year. The size of the schools in ECSDM ranges from 435 to 2,048 students. This district serves primarily Hispanic and African American populations (46.0% and 27.0%, respectively), with more than half of students (64.0%) identified as

[1] Georgia and New York were chosen because both states have a fairly extensive literacy assessment system.

"economically disadvantaged" by the New York Department of Education. Seven K-2 teachers trained in LLI and 218 K-2 students eligible for LLI in ECSDM participated in this study.

Methods

The present study of the LLI system employed a randomized controlled trial, mixed-methods design, which includes both quantitative and qualitative data and allows students to be randomly selected for the treatment (i.e., LLI in the first semester) or control (i.e., LLI in the second semester, if needed) condition. A matched-pair design was also utilized to ensure equivalency between treatment and control groups, and pre-post comparisons of student achievement in literacy were conducted. In addition, assessments of fidelity of LLI implementation included both independent observations and feedback from teachers and independent on-site researchers, and yielded both observational and self-reported survey data.

Multiple instruments were utilized in the evaluation, including two measures of reading achievement for evaluating students' progress in literacy; one observational tool for assessing teachers' LLI instructional practices; and two teacher surveys and focus groups to obtain teachers' and on-site researchers' feedback on LLI .

Procedure

The current study extended from March 2009 through June 2010. In the spring of 2009, three CREP researchers were responsible for ensuring that the districts understood and agreed to participate in the study while implementing LLI as intended by the developers. CREP researchers provided on-site orientation to the project and trained school coordinators and on-site researchers in each district to assist with data collection. At the beginning of the 2009-2010 school year, each district provided CREP with a list of first and second grade students that they had identified as eligible for LLI using their own selection criteria and whose parents had provided consent to participate in the study. Pre-testing of these students with the LLI Benchmarks and DIBELS began during the first three weeks of school. Subsequently, CREP conducted the randomization of the matched pairs of first and second graders based on demographic characteristics (i.e., gender, ethnicity, ELL status, special education status, and free/reduced lunch status) and pre-test LLI benchmark scores of instructional reading level. Students in the treatment group were then placed in LLI groups by LLI teachers, and the planned 90 days of LLI instruction for first and second graders began. Control group students did not receive LLI until the first and second grade evaluation period ended, and neither treatment nor control students received any additional pull-out literacy interventions during the study period.

Once at the beginning of the study period and once at the end, on-site researchers used the LLIOT to conduct random observations of each first and second grade LLI group. Post-tests with the LLI Benchmarks and DIBELS for the first and second grade students were completed at the conclusion of LLI in February for TCS and March for ECSDM. The LLI teachers and first and second grade classroom teachers with students in the study also completed an online survey regarding LLI or the school's core literacy program, as applicable, at this time. After CREP researchers conducted mid-year follow-up visits in each district, the entire procedure was repeated for kindergarten students, who began LLI in February (TCS) and April (ECSDM) and concluded in May (TCS) and June (ECSDM). Finally, CREP researchers visited each district at the end of the school year to address any remaining issues related to the study and to conduct structured focus groups with LLI teachers and on-site researchers.

Results

Student Achievement: Fountas & Pinnell LLI Benchmarks and DIBELS

Kindergarten LLI Benchmarks

On average after 38 days of LLI instruction, kindergartners who received LLI achieved a mean gain of 1.56 benchmark levels as compared to 0.78 benchmark levels for kindergartners who did not receive LLI. Also, kindergartners in LLI started, on average, below grade level in benchmark testing (i.e., pre-A = 0) but finished at a level between A and B, whereas their counterparts in the control group started near pre-A and finished around Level A. Thus, kindergartners in LLI finished the school year close to grade level in literacy (i.e., end-of-year kindergarten grade level goal = Level C). Also of note, English Language Learner (ELL), African American, and Hispanic students in LLI exceeded those in the control group. ELL students in LLI achieved a mean gain of about 1 ½ benchmark levels (M = 1.55) compared to a ½ benchmark level (M = 0.50) for ELL students not in LLI. African American LLI students also gained about 1 ½ benchmark levels (M = 1.44) while those in the control group only gained less than a benchmark level (M = 0.79). Finally, Hispanic students in LLI made the most gains—almost 2 benchmark levels (M = 1.76)—versus their counterparts in the control group who gained less than a benchmark level (M = 0.70). Also, all three subgroups finished closer to grade level (i.e., Level C) than their counterparts who finished around Level A or below.

Kindergarten DIBELS

Overall, fewer significant gains were seen with the DIBELS outcomes. However, kindergartners in LLI significantly exceeded those who were not in LLI on nonsense word fluency (NWF) (M = 10.64% and M = 6.88%, respectively). Also, for phoneme segmentation fluency (PSF), ELL students in the treatment group (M = 46.72%) outperformed ELL students in the control group (M = 23.96%), *as well as* non-ELL students in both the treatment and control groups (M = 23.24% and 24.24%, respectively). Thus, kindergartners who participated in LLI showed more significant gains on subtests of the DIBELS as compared to those who did not have LLI.

1st Grade LLI Benchmarks

On average after 73 days of LLI instruction, 1st graders who received LLI achieved a mean gain of 4.46 benchmark levels as compared to 2.63 benchmark levels for 1st graders who did not receive LLI. Also, 1st graders in LLI generally started below grade level in benchmark testing (i.e., A = 1) but finished at a level between E and F, whereas their counterparts in the control group started near Level A and finished around Level D. Thus, 1st graders in LLI finished their LLI sessions at the grade level mid-year goal in literacy (i.e., mid-year grade level goal for 1st grade = Levels E/F), while the control group students were still slightly behind. Also of note, African American and Hispanic students in LLI exceeded those in the control group. African American LLI students made the most gains—they gained about 5 ½ benchmark levels (M = 5.20) while those in the control group only gained about 2 ½ benchmark levels (M = 2.60). Finally, Hispanic students in LLI also made significant gains—about 4 benchmark levels (M = 4.18)—versus their counterparts in the control group who gained about 2 ½ benchmark levels (M = 2.57). Also, both subgroups finished at the grade level goal (i.e., Level E/F) compared to their counterparts in the control group who finished close to Level D. Of importance to note, the finding for African American 1st graders in LLI appears particularly robust and educationally significant. These LLI students finished the highest out of all subgroups as well as the aggregate—close to Level G—versus all others who finished between Levels C to F.

1st grade DIBELS

Overall, similar significant differences between treatment and control groups were seen with the 1st grade DIBELS outcomes. 1st graders in LLI significantly exceeded those who were not in LLI on nonsense word fluency (NWF) (M = 22.00% and M = 17.00%, respectively). Also, for NWF, Hispanic students in the treatment group (M = 19.00%) outperformed their counterparts in the control group (M = 17.00%). Additionally, 1st graders who received LLI performed better than their counterparts on Oral Reading Fluency (ORF) (M = 14.00% and M = 11.00%, respectively), as well as on Letter Naming Fluency (LNF) (M = 17.00% and M = 11.00%, respectively). Thus, 1st graders who participated in LLI showed more significant gains on subtests of the DIBELS as compared to those who did not have LLI.

2nd Grade LLI Benchmarks

On average after 73 days of LLI instruction, 2nd graders who received LLI achieved a mean gain of 4.64 benchmark levels as compared to 2.99 benchmark levels for 2nd graders who did not receive LLI. Also, 2nd graders in LLI started, on average, below grade level in benchmark testing (i.e., E = 5) but finished at Level J, whereas their counterparts in the control group started closer to Level F but only finished around Level I. Thus, 2nd graders in LLI finished the school year close to the grade level mid-year goal in literacy (i.e., mid-year grade level goal for 2nd grade = Level J/K). Also of note, a robust overall effect was found for students with a special education designation who received LLI. These students in the treatment group started around Level C and finished close to Level H, while their counterparts in the control group started at Level D and finished around Level F. Also, regarding ethnicity subgroups, White students in LLI finished above their counterparts in the control group, gaining about 5 benchmark levels (M = 5.05) compared to about 3 benchmark levels (M = 3.14) in the control group. Additionally, African American and Hispanic students in LLI exceeded their counterparts in the control group. Of particular educational significance, African American LLI students finished at the highest level compared to all others—just above Level I; however, this was closely followed by the Hispanic LLI students who also finished slightly above Level I on average. The African American students in the treatment group gained about 4 ½ benchmark levels (M = 4.46), while those in the control group only gained about 2 ½ benchmark levels (M = 2.67). Finally, Hispanic students in LLI gained more than African American students in LLI (M = 4.53 and M = 4.46, respectively), while Hispanic students in the control group only gained about 3 benchmark levels.

2nd Grade DIBELS

Overall, no significant differences were found between treatment and control groups for 2nd grade on either DIBELS subtest that was administered as intended for 2nd graders (i.e., Nonsense Word Fluency and Oral Reading Fluency). While unexpected, this result may simply indicate that the 2nd grade DIBELS measures were not sufficiently in alignment with the 2nd grade LLI curriculum or benchmarks to detect small effects, or changes, in DIBELS scores. However, it is also plausible that the lack of an overall effect may be due to district-level differences in these scores. One district appears to have made significant gains on the 2nd grade DIBELS tests compared to the other, but taken together, no overall effects were able to be seen (i.e., a wash-out effect from averaging across both districts' scores).

Observations: Leveled Literacy Observation Tool (LLIOT)

The results from the LLIOT revealed that 5 of the 10 LLI lesson components were rated "Acceptable" or "Excellent" over 90% of the time, indicating a high level of implementation fidelity across both districts. The highest rated lesson components (i.e., those demonstrating the highest degree of implementation fidelity) included phonics/word work, reading a new book, and rereading. The lowest rated lesson components (i.e., those demonstrating the lowest degree of implementation fidelity) included classroom and home connections. Teachers were also rated highly on their use of literacy instructional strategies, such as modeling and encouraging fluent oral reading and appropriate reading strategies and assisting students in problem-solving. Further, in the majority of observed lessons, instructional materials were readily available; the lesson was well-organized; and students were engaged and attentive. Additionally, the majority of observed groups had 3 students and lasted approximately 30 minutes, which was consistent with LLI's design. Overall, observers perceived that the lesson was delivered as designed 96.3% of the time.

The LLIOT was conducted at both the beginning and the end of LLI for each of the observed groups in order to measure changes in implementation over time. For the 25 observed kindergarten groups, there were no significant differences on any of the 3 LLIOT subscales (Quality of LLI Implementation, Literacy Instructional Strategies, and Learning Environment) from the first observation to the second. For both the 25 observed first grade groups and the 33 observed second grade groups, only scores on the Learning Environment subscale improved significantly from pre-test to post-test. For each subscale at each grade level, the average rating was between "Acceptable" and "Excellent" at both time points.

Teacher Surveys: LLITQ & CTLIQ

Overall, on the Leveled Literacy Intervention Teacher Questionnaire (LLITQ), LLI teachers were most likely to agree that they understood the goals and implementation procedures for LLI, that LLI positively impacts student literacy achievement, and that their districts and other teachers within their schools were supportive of LLI. LLI teachers also reported a positive impact of LLI on their reading instruction, particularly their understanding of the role of comprehension and phonics/phonemic awareness in the reading process and the relationship of leveled texts to successful reading. LLI teachers were least likely to agree that the parents of their LLI students participated in home literacy activities with their children, that their schools had sufficient faculty and staff to provide LLI to all students who needed it, and that LLI helped their students with special needs and ELL students. All of the surveyed teachers agreed that their school should continue using the LLI system.

In terms of the regular classroom literacy instruction provided to both treatment and control students in the study, results from the Classroom Teacher Literacy Instruction Questionnaire (CTLIQ) revealed that the K-2 classroom teachers were most likely to provide individual or small-group reading instruction, integrate vocabulary and comprehension into their literacy instruction, and utilize high-quality literature to read to students and engage them in interactive discussions about the text. Teachers were least likely to report utilizing whole-class reading instruction and assigning home literacy activities for students to complete with parents. Overall, the classroom teachers were most likely to agree that they understood the goals of their literacy program, that it was aligned with state and district reading/language arts standards, and that their faculty, staff, and administration believed that all students could learn to read and write. Similar to the LLI teachers, classroom teachers were least likely to agree that the parents of their students participated in home literacy activities with their children,

that their schools had sufficient faculty and staff to fully implement their literacy program, and that their literacy program helped their students with special needs and ELL students. The majority of surveyed teachers agreed that their school should continue their current literacy program.

Focus Groups

Structured focus groups conducted with the LLI teachers in the study revealed that most of the LLI teachers liked LLI and felt that it was beneficial to their students. Some teachers felt that the system needed more work, and others felt that school-level variables (e.g., support, time, materials) needed improvement in order to implement LLI correctly. LLI teachers reported that the most frequently encountered logistical issue when implementing LLI was time and/or scheduling of LLI groups to coordinate with classroom teachers' schedules and complete lessons during the designated 30-minute timeframe. In terms of strengths, LLI teachers most frequently identified the instructional materials, particularly the books and take-home books. LLI teachers also liked the design (e.g., group size, lesson layout, guided format of lessons). When asked about areas of improvement for LLI , LLI teachers most frequently mentioned an inability to adequately complete a lesson in 30 minutes, an inconsistency of materials (e.g., the lesson did not "match" the written materials), and the fact that the system was too fast-paced for their lower-level students. LLI teachers also discussed problems with using the new online LLI Data Management System, including slowness and missing data, and recommended on-site training and additional resources for using the system.

Structured focus groups were also conducted with on-site researchers who completed the DIBELS assessments and LLIOT observations for the study. Based on their observations of LLI lessons, on-site researchers described the LLI teachers' group management skills as a strength of the LLI implementation during the current study. On-site researchers were also impressed with the well-organized, adaptable nature of LLI and its ability to build student confidence. When asked about areas of improvement for LLI , on-site researchers most frequently mentioned the length of the Reading Records, the difficulty of completing a lesson in 30 minutes, and the fact that the system was too fast-paced for slower learners. Overall, the on-site researchers in the focus groups were positively supportive of LLI, but they did caution that the system's effectiveness could be affected by the teacher's experience and level of LLI training.

Conclusions

1. What progress in literacy do students who receive LLI make compared to students who receive only regular classroom literacy instruction?

Across the three grade levels, the current study found that LLI positively impacts K-2 student literacy achievement in rural and suburban settings. Further, we determined that LLI is effective with ELL students, students with a special education designation, and minority students in both rural and suburban settings. Finally, the current study showed that LLI is effective with economically disadvantaged children in both rural and suburban settings.

This study found robust effects on the LLI Benchmarks across all grade levels for students who received LLI. Across the three grade levels, students in LLI achieved between 1 ½ benchmark levels up to almost 5 ½ benchmark levels, while students who did not receive LLI achieved between less than 1 benchmark level to 3 benchmark levels.

Further, these effects were particularly strong for various subgroups (e.g, ethnicity, special education or ELL status) within each grade level. For kindergarten, significant effects were found, compared to the control group, for African American students, Hispanic students, and ELL students on the LLI Benchmarks, with all three subgroups finishing closer to grade level (i.e., Level B) than their counterparts who finished at or below Level A. First grade African American and Hispanic students in the treatment group also showed more gains than their counterparts in the control group. In second grade, strong, educationally meaningful effects were found for African American and Hispanic LLI students. Second grade African American LLI students finished at the highest level overall, closely followed by the Hispanic LLI students.

Additionally, effects found with the DIBELS measures of reading fluency provided corroboration of the results with the LLI Benchmarks. In kindergarten, students in LLI showed significant gains on subtests of the DIBELS as compared to those who did not have LLI. In particular, for phoneme segmentation fluency, ELL students in the treatment group outperformed ELL students in the control group, *as well as* non-ELL students in both the treatment and control groups. In 1st grade, LLI students significantly exceeded the control group on 3 of 4 subtests: nonsense word fluency, letter naming fluency, and oral reading fluency. Finally, on the nonsense word fluency subtest, 1st grade Hispanic students in the treatment group outperformed their counterparts in the control group.

Taken together, all of the student achievement results provide strong evidence that students who are eligible for and participate in LLI make significant progress in literacy compared to students who are eligible to receive LLI and only receive regular classroom literacy instruction.

2. Was LLI implemented with fidelity to the developers' model?

Across all observations, the observation results from the current study suggest that LLI was implemented with a high degree of fidelity to design across both districts. The majority of lesson components received high fidelity ratings in most of the observations that were conducted. Additionally, observation results revealed that LLI implementation was consistent across the year, with high fidelity scores received at both time points when the observations were conducted. Finally, although students received, on average, less than the model's recommended number of instructional days, students in all three grade levels made significant progress in their literacy achievement. This finding suggests that LLI can still be effective during a relatively short timeframe, which may be valuable to districts with a large number of students to serve or limited time in which to implement early literacy interventions.

3. What were LLI teachers' perceptions of LLI and its impact on their students' literacy?

Overall, the LLI teachers in the current study supported LLI and believed that it had a positive impact on their students' literacy achievement and attitudes toward literacy. LLI teachers indicated that they had a good understanding of the system; received support in implementing LLI from their district, school administration, and other school staff; and perceived a positive impact of LLI on their reading instruction. LLI teachers were particularly impressed with the system's leveled texts as well as the small-group format and guided lesson structure; however, many LLI teachers felt that the lessons could not be completed in 30 minutes, that the system was too fast-paced for their lower-level students, and that there were some inconsistencies in the materials. Finally, in addition to the LLI teachers, a small number of classroom teachers with students in the current study provided feedback on their perceptions of the LLI system. Most of these teachers were positive about the system and noticed that their students'

literacy in the classroom improved after receiving LLI, with one classroom teacher even commenting, "…I believe that children that struggle would give up hope in the realm of reading without the LLI program."

Recommendations

Altogether, the results from this evaluation allow us to conclude that LLI positively impacts students' literacy skills. These results also suggest that continued implementation of LLI would be beneficial in both Tift County Schools and the Enlarged City School District of Middletown. While the long-term impacts of LLI have yet to be determined, the positive results found in this evaluation suggest that additional benefits may be seen with the continuation of LLI. This evaluation provided a randomized controlled trial and efficacy study for the LLI system as well as offered an opportunity for research-based recommendations that may enhance the system, future research, and ultimately student achievement. From this evaluation, CREP proposes the following recommendations with regard to LLI and its implementation in schools:

- When possible, schools should begin kindergarten instruction in LLI as soon as possible in order to provide the recommended amount of instruction (i.e., 14 weeks) for kindergarten students.

- Professional development for building principals and central office supervisory staff, although not measured in this study, surfaced as being critical to the implementation.

- Likewise, regular classroom teacher involvement and professional development to familiarize them with LLI and its features also appears to influence the quality of implementation.

- LLI teacher professional development should be ongoing with at least a refresher training to supplement and resolve any district-specific issues.

- Providing scenarios or examples of how prior adopters have developed schedules that allow for full implementation of the 30-minutes-a-day, five-days-a-week instructional pattern would be helpful to school districts who are new adopters of LLI.

- Suggestions and recommendations of how LLI teachers might plan and organize their LLI sessions so they can accomplish the instructional goals in a typical 30-minute session would benefit prior and new adopters of LLI.

- Additional suggestions from the authors about how best to instruct LLI groups whose members are not at the same level or who have members progressing at a slower rate would be helpful.

- Providing some type of video for parents of the LLI students could not only explain the system but could provide clips of how they should be working with their child. This is particularly important for the parents of ELL children and the parents of economically disadvantaged children.

- A careful review of all materials and resources is recommended to ensure consistency and accuracy throughout the system.

- There is a great need to conduct a similar study in at least one major urban district.

- Future research of LLI should include longitudinal tracking of student reading achievement to look at the long-term impact of LLI beyond one school year.

- The LLI benchmarking system would benefit from additional systematic comparisons with other nationally recognized literacy assessments.

Introduction

This report summarizes evaluation results for an efficacy study of the Leveled Literacy Intervention system (LLI) implemented in Tift County Schools (TCS) in Georgia and the Enlarged City School District of Middletown (ECSDM) in New York during the 2009-2010 school year. Both school districts adopted the targeted, small-group implementation model of LLI in their schools with support from Heinemann consultants providing LLI professional development. This report focuses on the implementation and impact of this model during the first full school year of the system in these schools.

Developed by authors Irene C. Fountas and Gay Su Pinnell (2009) and published by Heinemann, LLI is a short-term, small-group, supplemental literacy intervention system designed for students in kindergarten through second grade who struggle with literacy. The goal of LLI is to provide intensive support to help these early learners quickly achieve grade-level competency. The LLI materials are based around a series of "leveled" texts (i.e., texts of progressing difficulty) with difficulty measured by the Fountas & Pinnell Text Level Gradient™, A-Z (Fountas & Pinnell, 2007). The system emphasizes systematic and explicit instruction in phonological awareness, phonics, fluency, comprehension, and the expansion of oral language skills, including vocabulary.

Heinemann consultants provided professional development sessions for teachers regarding the LLI materials and instructional strategies. The teachers also received professional development training on the LLI online data management system, used to track student progress and attendance. The evaluation was designed to examine the extent to which participation in LLI influenced student literacy achievement and teachers' instructional practices regarding literacy. Additionally, this study was designed to determine the strengths and weaknesses of LLI according to relevant stakeholders.

The work reported here was conducted by the Center for Research in Educational Policy (CREP), a State of Tennessee Center of Excellence, located at the University of Memphis. The Center's 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, the Center seeks to accomplish its mission through a series of investigations conducted by Center personnel, college and university faculty, and graduate students.

Theoretical Framework

Research suggests that children with poor early reading skills continue to struggle with reading and writing in the later grades and are more likely to drop out of school (Alexander, Entwisle, & Horsey, 1997; Juel, 1988; Tabors, Snow, & Dickinson, 2001). However, there is evidence that quality early intervention programs can prevent the development of long-term reading deficiencies (Heibert & Taylor, 1994; Wanzek & Vaughn, 2007). Previous studies by Harrison, Peterman, Grehan, Ross, Dexter, and Inan (2008) and Peterman, Grehan, Ross, Gallagher, and Dexter (2009) showed that K-2 students enrolled in LLI made significant gains on the Gates-Mac Ginitie Reading Test, with 25 to 44% of students reading at or above average by the end of the study. The LLI system has its roots in the theoretical and empirical work of Marie Clay (1991) and of Fountas and Pinnell (1996, 2006), and its lesson design draws from empirical research on reading acquisition and reading difficulties, language learning, and student

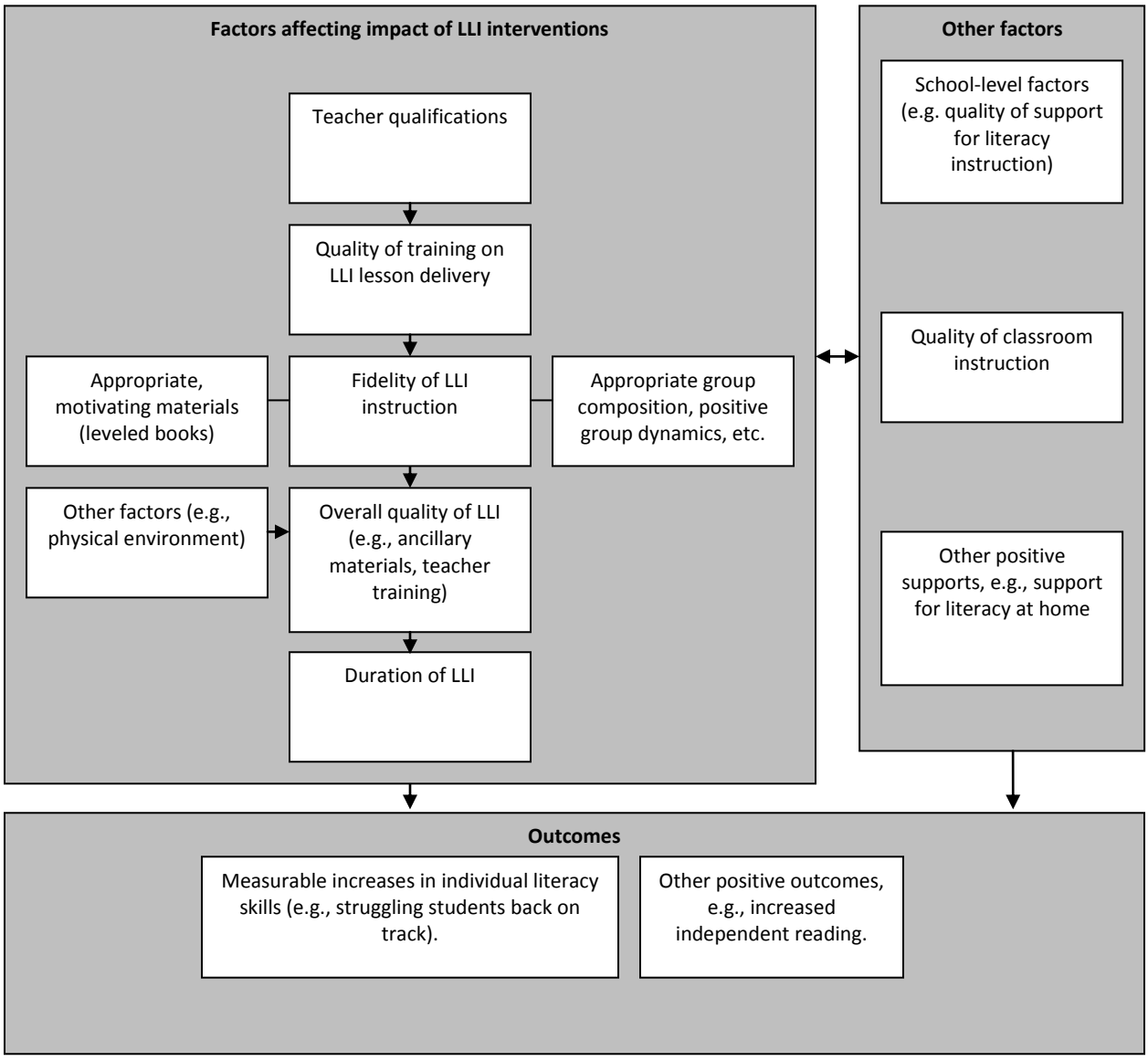
motivation (e.g., Armbruster, Lehr, & Osborn, 2001; National Institute of Child Health and Human Development, 2000a; National Institute of Child Health and Human Development, 2000b).

The current study expanded on these findings by utilizing a multi-site, randomized experimental design to examine whether students in LLI achieved greater gains in literacy than students receiving classroom literacy instruction alone.

Theory of Action

Figure 1, below, represents our preliminary identification of some of the key factors that impact the quality of LLI and its overall impact on student learning. Generally, the model states that measurable increases in student literacy growth and other positive outcomes will result upon completion of a certain number of intervention sessions as well as from a combination of factors, including those directly related to the intervention itself—*and* other non-LLI factors such as the quality of the students' regular classroom instruction and support they receive for literacy at home.

Figure 1: LLI Theory of Action



LLI Factors

Factors that we propose that may directly affect the quality of LLI include: teacher qualifications and skills; the quality of training; the level of materials matched appropriately to students' reading level and progress (e.g., teachers select appropriate sequence of leveled books at the students' reading level); the overall quality and fidelity of LLI instruction; composition of the student group (e.g., students at more or less the same reading level or combinations that include students with special needs and ELL students); and other factors such as the learning environment and the duration of the intervention. We discuss each of these factors in turn, along with how these factors may be measured.

Teacher Qualifications and Training

A cluster of teacher factors may affect the overall quality of LLI. For example, teachers who already have a certain level of experience and skill in delivering literacy interventions may be more likely to benefit from LLI training and more likely to make good use of the materials than teachers who are less experienced. It was, therefore, important to have background information on the teachers providing the intervention—including years of experience, degrees attained, and other relevant training and job experience. Finally, the professional development provided to LLI teachers, subsequent to their selection, is critical to LLI implementation.

Appropriate Selection of LLI Materials

The LLI system depends heavily on the use of leveled, high-interest texts that are selected after assessing students with the Fountas & Pinnell benchmarks in order to determine each student's beginning instructional reading level and independent reading level. As the intervention progresses, teachers select the progress or sequence of the leveled texts that students read. Therefore, it was important to evaluate the match between the leveled texts used for instruction, at the beginning and throughout the LII intervention. It was also important to assess students' degree of engagement in using the materials. At a minimum, this implies a careful log of the books used along with data on each student's instructional reading level. Careful observations of the interventions themselves, along with teacher surveys and logs from the LLI data management system, were used to measure students' achievement and the level of student engagement during lesson instruction.

Fidelity of LLI

High-fidelity implementation of LLI depends in part on the amount and quality of professional training provided, the support of school administrators, and a commitment from the teachers selected to be LLI interventionists. We also wanted to know how much training had been provided—as well as the quality and relevance of the training from the teachers' perspective. Relevant teacher demographics and perceptions were obtained from the participating teachers. The developers provided details about the professional development that was provided to each district.

Group Composition and Behavior

Since LLI is a small-group intervention, it was important to know about the individual characteristics of each group. For example, if students have different needs (e.g., students can be struggling at the same instructional reading level for different reasons), it could be difficult for a teacher to provide instruction that meets the needs of the group as a whole. Also, if one student is unmotivated

or disruptive, this could presumably impact the social dynamics of the group—and ultimately affect the success of the intervention. It was also important to know whether some or all of the students in the group had been identified as students with special needs or ELL students. With this in mind, we did ascertain each group's demographics and characteristics (i.e., literacy level).

Other Intervention Factors

The impact of LLI is directly affected by other factors, such as the duration of the intervention. We utilized part of the LLI data management system to obtain a record of the actual number of days of intervention for each group (14 weeks is recommended for kindergarten and 18 weeks is recommended for first and second grades). In addition, data was collected relative to adherence to the recommended 30-minutes-a-day, 5-days-a-week instructional cycle. Finally, we conducted structured focus groups with on-site researchers and LLI teachers to gather additional qualitative data related to instructional time, student absences and mobility, materials, and several other tangential factors.

Non-LLI Factors

Non-LLI factors include school-level variables, such as the overall support for literacy in the school and the quality of instruction in the child's regular classroom.

School-Level Variables

A full understanding of how LLI works in a particular context—and why it may be more or less successful from one school to another—will be usefully informed by understanding school-level factors, such as overall support for literacy in the school (e.g., literacy may receive more emphasis and resources in some schools than in others) and school-level attention to the needs of struggling students. Certain schools, in other words, may provide contexts that tend to promote a high-quality implementation of LLI. In the current study, we measured these factors through surveys of both LLI teachers and regular classroom teachers.

Quality of Regular Classroom Literacy Instruction

Students receiving LLI in the current study were also receiving literacy instruction from their regular classroom teacher, and some portion of any measured gains in literacy skills over the period of the study may be attributable to the quality of literacy instruction that the children received in their regular classroom. Students who are receiving high-quality literacy instruction in the classroom *and* high-quality intervention are more likely to show progress than students who receive the same quality of intervention but lower-quality classroom instruction. We used a teacher survey regarding the school's literacy program as a measure of the nature of regular classroom instruction that the intervention students received.

Research Questions

The purpose of this study was threefold: (1) to determine the efficacy of the Leveled Literacy Intervention system (LLI) in increasing reading achievement for K-2 students; (2) to examine the implementation fidelity of LLI; and (3) to determine perceptions of LLI according to relevant stakeholders. This study focused on two U.S. school districts and comprised 427 K-2 students who were

matched demographically and randomly assigned to treatment and control groups. The evaluation used a mixed-methods design to address the following key research questions:

1. What progress in literacy do students who receive LLI make compared to students who receive only regular classroom literacy instruction?

2. Was LLI implemented with fidelity to the developers' model?

3. What were LLI teachers' perceptions of LLI and its impact on their students' literacy?

Methods

The present study of the LLI system employed a randomized controlled trial, mixed-methods design, including both quantitative and qualitative data. A matched-pair design was also utilized to ensure 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 both independent observations and feedback from teachers and independent on-site researchers, yielded both observational and self-reported survey data.

Multiple instruments were utilized in the evaluation, including two measures of reading achievement for evaluating students' progress in literacy; one observational tool for assessing teachers' LLI instructional practices; and two teacher surveys and focus groups to obtain teachers' and on-site researchers' feedback on LLI . Details of each instrument will be discussed later in this section.

System Description: Leveled Literacy Intervention (LLI)

The Leveled Literacy Intervention system (Fountas & Pinnell, 2009) is a short-term, intensive, small-group intervention designed for children in kindergarten through second grade who are having difficulty learning early reading and writing skills. The goal of the system is to accelerate these children's progress in order to bring their skills up to grade level so their early literacy difficulties do not become long-term deficits. The system is appropriate for struggling regular education students and students with special needs, and there are minor modifications for English language learners (ELL students).

Children enrolled in LLI meet in small groups (ideally three students) for daily 30-minute lessons, and the intervention lasts a maximum of 18 weeks, depending on the progress of the individual child. According to developers (Fountas & Pinnell, 2008), LLI emphasizes the development of oral language skills as a foundation for reading and the five components of reading instruction identified by the National Reading Panel (National Institute of Child Health and Human Development, 2000a): phonological awareness and phonics, fluency, vocabulary, and comprehension. Phonics instruction is systematic, explicit, and follows a prescribed sequence of sound-letter relationships and spelling patterns. Additionally, reading comprehension skills are taught through intensive interactions between the teacher and the students and amongst students. LLI also is designed to develop students' motivation and interest in reading and writing.

An underlying premise of LLI is that children benefit from experience with texts that they can read without difficulty at their "independent level," as well as with more challenging texts written at their "instructional level" (Fountas & Pinnell, 2008). The LLI system provides students with both kinds of

reading experiences, alternating between easier texts and more challenging ones. Easier texts build fluency and give students success at reading that builds confidence and positive self-esteem. More challenging texts, which students read with scaffolding and support from the LLI teacher, give children the opportunity to develop more sophisticated reading skills. LLI materials specify concepts that teachers can emphasize when discussing each book in the sequence. Other key ideas underlying the design of LLI are the following:

- Struggling children learn best when lessons follow a predictable sequence. All LLI lessons have the same basic structure, allowing children to focus most of their processing attention on reading, writing, phonics, and word study activities.

- Children who are struggling with reading and writing need to learn fast, automatic processing of oral and written language. For this reason, LLI lessons are designed to be fast-paced, with a specified set of literacy activities for each day of the intervention. The fast pace promotes rapid processing and keeps children engaged in the lessons and motivated to participate in the literacy activities and discussion.

- Literacy interventions should be linked to classroom instruction and the home environment. Children take LLI books home to read aloud to their parents, along with simple homework assignments, and they also may take books back to the classroom.

- A system of ongoing formative assessments conducted during the 18 weeks gives teachers information about student learning that can inform their instructional decision-making.

Literacy teachers selected to be LLI teachers receive eight days of professional development focused on how to implement the LLI instructional program. They also receive the necessary LLI materials and a detailed teaching guide. Additional professional development is provided throughout implementation, including training in how best to facilitate comprehension skills through teacher-student and student-student interactions.

In addition to the professional development regarding LLI materials and instructional strategies, TCS and ECSDM educators also received training on the LLI online data management system, used to track student progress and attendance. Further, for the purposes of the study, the two school districts voluntarily agreed to provide the LLI system as specifically designed by the developers. A strict implementation plan was utilized, which included the following guidelines:

- No additional pull-out literacy interventions for either treatment or control students for the duration of the study

- The maximum number of instructional days (i.e., each district attempted to provide 90 days of LLI instruction to first and second graders and 70 days to kindergarteners, according to the recommendations of developers)

- Three students per group

- Consistent LLI completion across the district (i.e., all students in the study within each district would begin and end LLI at the same time)

Setting and Population of Participants

Five elementary schools in Tift County Schools (TCS) in Tifton, Georgia, and four elementary schools in the Enlarged City School District of Middletown (ECSDM) in Middletown, New York, volunteered to participate in the study.[2] TCS is a rural school district in a small town located approximately 181 miles south of Atlanta, Georgia, that served 7,551 students during the 2008-2009 school year. Most of the schools in TCS are small and serve primarily White and African American populations (48.0% and 35.0%, respectively), with more than half of students (65.0%) identified as "economically disadvantaged" by the Georgia Department of Education's free and reduced lunch status. Twenty-one K-2 teachers trained in LLI and 209 K-2 students eligible for LLI in TCS participated in this study.

ECSDM is a suburban school district in a small city located approximately 72 miles northwest of New York City, New York, that served 6,764 students during the 2008-2009 school year. The size of the schools in ECSDM ranges from 435 to 2,048 students. This district serves primarily Hispanic and African American populations (46.0% and 27.0%, respectively), with more than half of students (64.0%) identified as "economically disadvantaged" by the New York Department of Education's free and reduced lunch status. Seven K-2 teachers trained in LLI and 218 K-2 students eligible for LLI in ECSDM participated in this study. Table 1 summarizes the overall demographic characteristics of both districts.

Table 1: Demographic Overview of TCS and ECSDM Schools (PreK-12)

		School Wide Population		Student Population						
School District	Grade Levels	Students	Teachers	% Asian	% African American	% Hispanic	% White	% Economically Disadvantaged	% Students with Disabilities	% English Language Learners
Tift County	PK-12	7551	552	1.0	35.0	13.0	48.0	65.0	11.0	8.0
Middletown	PK-12	6764	478	2.0	27.0	46.0	25.0	64.0	6.9	12.0

Note: Demographic information for TCS obtained from 2008-09 School Report Card and http://nces.ed.gov/ccd/districtsearch/district_detail.asp?Search=1&details=+&InstName=tift&State=13&DistrictType=1&DistrictType=2&DistrictType=3&DistrictType=4&DistrictType=5&DistrictType=6&DistrictType=7&NumOfStudentsRange=more&NumOfSchoolsRange=more&ID2=1304980
Demographic information for ECSDM obtained from 2008-09 School Report and http://nces.ed.gov/ccd/districtsearch/district_detail.asp?Search=1&City=+middletown&State=36&DistrictType=1&DistrictType=2&DistrictType=3&DistrictType=4&DistrictType=5&DistrictType=6&DistrictType=7&NumOfStudentsRange=more&NumOfSchoolsRange=more&ID2=3619320&details=

Teacher Demographics

A total of 28 LLI teachers and 125 classroom teachers across both districts participated in this study. According to data obtained from a survey of all participating LLI teachers, the majority of LLI teachers in the study had been teaching in their current school (84.1%) or any school (93.2%) for 6 or more years. Most LLI teachers had also completed a Master's degree or beyond (65.9%). LLI teachers were all female, 97.7% White, and almost all held their professional teaching certification (95.5%). Additionally, almost all of the LLI teachers in the study (93.2%) had completed the LLI professional development. Overall, these teachers had a solid background of teaching experience at their current

[2] Georgia and New York were chosen because both states have a fairly extensive literacy assessment system.

school and teaching in general. Around two-thirds of them had pursued advanced degrees and continuing education in their field. Taken together, they appear to have been well positioned to receive and implement a new 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 = 28)

Item	Percent Responded
Years of teaching experience at current school	
5 years or less	15.9
6-10 years	40.9
11 or more years	43.2
Years of teaching experience at any school	
5 years or less	6.8
6-10 years	18.2
11 or more years	75.0
Highest level of education completed	
Bachelor's Degree	34.1
Master's Degree	31.8
Master's plus 30 hours, Education Specialist, or Doctoral Degree	34.1
Ethnicity	
Asian or Pacific Islander, American Indian or Alaskan Native, or Multi-racial/other	0.0
African-American/ Black	2.3
Hispanic	0.0
White, not of Hispanic origin	97.7
Gender	
Male	0.0
Female	100.0
Age group	
29 years or less	4.5
30-39 years	22.7
40-49 years	29.5
50-59 years	31.8
60 years or older	11.4
Level of LLI training	
Completed training	93.2
Partially trained	2.3
None	2.3
Teacher certification level	
Paraprofessional	0.0
Alternative certificate	0.0
Initial/apprentice certificate	4.5
Regular/professional certificate	95.5

Note: Item percentages may not total 100% because of missing input from some respondents.

According to data obtained from a survey of 89 of the 125 participating classroom teachers, the classroom teachers in the current study were fairly evenly distributed across K-2 grade levels (31.5%, 36.0%, and 32.6%, respectively). The majority had been teaching in their current school (67.4%) or any school (77.5%) for 6 or more years. Most K-2 classroom teachers had also completed a Master's degree or beyond (74.1%). K-2 classroom teachers were 96.6% female, 94.4% White, and 98.9% held their professional teaching certification. Overall, the participating classroom teachers generally had a good level of experience with their current school and teaching in general, and nearly three quarters of them had pursued advanced degrees and continuing education in their field. Table 3 summarizes the

demographic characteristics of the classroom teachers in the study, as reported on the classroom teacher survey.

Table 3: Demographic Characteristics of Participating K-2 Classroom Teachers (n = 89)

Item	Percent Responded
Grade level	
K	31.5
1	36.0
2	32.6
Years of teaching experience at current school	
5 years or less	32.6
6-10 years	34.8
11 or more years	32.6
Years of teaching experience at any school	
5 years or less	21.3
6-10 years	24.7
11 or more years	52.8
Highest level of education completed	
Bachelor's Degree	24.7
Master's Degree	60.7
Master's plus 30 hours, Education Specialist, or Doctoral Degree	13.4
Ethnicity	
Asian or Pacific Islander, American Indian or Alaskan Native, or Multi-racial/other	1.1
African-American/ Black	1.1
Hispanic	1.1
White, not of Hispanic origin	94.4
Gender	
Male	3.4
Female	96.6
Age group	
29 years or less	18.0
30-39 years	38.2
40-49 years	20.2
50-59 years	19.1
60 years or older	4.5
Teacher certification level	
Paraprofessional	0.0
Alternative certificate	0.0
Initial/apprentice certificate	1.1
Regular/professional certificate	98.9

Note: Item percentages may not total 100% because of missing input from some respondents.

Student Demographics

Across the five participating schools in TCS and the four participating schools in ECSDM, there were a total of 427 students who participated in this study. Of these students, 146 were in kindergarten, 130 were in first grade, and 151 were in second grade. A total of 222 students comprised the randomly assigned treatment group for the study, while 205 students made up the control group; the slight discrepancy in group size is attributable to student attrition that occurred after randomization was completed. On average across both districts' participating schools, 37.0% of students in the sample were Hispanic, 33.5% were African American, and 28.5% were White. The majority of participating students (84.5%) qualified for free or reduced price lunch, 13.5% were English Language Learner (ELL)

students, and 8.5% were classified as eligible for special education services. Table 4 summarizes the participating K-2 students across both districts.

Table 4: Demographic Overview of Participating K-2 Students (n = 427)

School District	Grade Levels	Students	% African American	% Hispanic	% White	% Other/Mixed Ethnicity	% Economically Disadvantaged	% ELL	% SpEd
Tift County	K-2	209	39.0	31.0	29.0	1.0	89.0	24.0	12.0
Middletown	K-2	218	28.0	43.0	28.0	1.0	80.0	3.0*	5.0*

Note: Demographic information obtained from each school district's records; ELL = English Language Learners; SpEd = students with a special education designation

*Middletown limited the number of ELL students and students with special education status who could participate in the study due to sheltered classrooms.

Instrumentation

Both quantitative and qualitative data were collected in this evaluation. CREP researchers used two measures of reading achievement for evaluating students' progress in literacy: the Fountas & Pinnell Benchmark Assessment System and the Dynamic Indicators of Basic Early Literacy Skills (DIBELS). One observational tool, the Leveled Literacy Intervention Observation Tool (LLIOT), was used to evaluate LLI literacy practices and instructional strategies in the classroom. Two teacher surveys, the Leveled Literacy Intervention Questionnaire (LLITQ) and the Classroom Teacher Literacy Instruction Questionnaire (CLITQ), were also used to ascertain teachers' feedback on LLI and classroom literacy instruction. Additionally, structured focus groups were conducted with LLI teachers and on-site researchers to gather additional qualitative feedback regarding LLI. Details of each instrument are discussed below.

Student Literacy Achievement

Fountas & Pinnell Benchmark Assessment System (LLI Benchmarks)

The Fountas & Pinnell Benchmark Assessment System was used to measure the following literacy skills: phonemic awareness, letter-sound relationships (decoding), vocabulary, comprehension, fluency, and writing. Both treatment and control students in the study were tested by LLI teachers at the beginning and the end of LLI. This data was 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 is an individually administered assessment tool designed by the developers of LLI to reliably place K-2 students on the Fountas & Pinnell Text Level Gradient™, A-Z (Fountas & Pinnell, 2007), an A-Z gradient of text difficulty. LLI is comprised of three systems: Levels A-C are in the Orange System; Levels A-J are in the Green System; and Levels C-N are in the Blue System. The Orange System is generally used in Kindergarten; the Green System in Grade 1; and the Blue System in Grade 2. The goal of the LLI system is to bring children up to their current grade level in reading, starting from the earliest Level A (mid-kindergarten) to Level N (early third grade).[3]

[3] Product description from the Heinemann LLI Field Study Request for Proposal

Dynamic Indicators of Basic Early Literacy Skills (DIBELS)

DIBELS, developed by the Early Childhood Research Institute on Measuring Growth and Development at the University of Oregon, is a tool for early identification of children with potential literacy problems and an assessment of response to instruction. The DIBELS assessment is designed to enable educators to modify their approach if a student is not on course to achieve reading goals. The Institute reports that it has validated the instrument's ability to predict outcomes and has tested its reliability with young children across the country. The measures were developed based upon the essential early literacy domains discussed in both the National Reading Panel (National Institute of Child Health and Human Development, 2000a) and National Research Council (1998) reports to assess student development of phonological awareness, alphabetic understanding, and automaticity and fluency with the code.

DIBELS consists of seven subtests, most of which can be used for either benchmark or progress-monitoring assessments. Different subtests are administered depending on the time of year and the grade. Benchmark assessments are given to all children in a grade three times a year, while progress-monitoring assessments are used electively. Administrating each subtest should take from five to seven minutes. The seven DIBELS subtests[4] are:

- Initial Sound Fluency (Pre-Kindergarten through Mid-Kindergarten) – This subtest measures the child's ability to identify, isolate, and pronounce the first sound of an orally presented word.

- Letter Naming Fluency (Kindergarten through Beginning of Grade 1) – This subtest asks students to name as many letters, both uppercase and lowercase randomly mixed, as they can in one minute.

- Phoneme Segmentation Fluency (Mid-Kindergarten through Grade 1) – This subtest is a direct measure of phoneme awareness. Students are asked to say the individual sounds that make up a word or syllable containing three or four phonemes.

- Nonsense Word Fluency (Mid-Kindergarten through Beginning of Grade 2) – This subtest measures a student's ability to link letters with sounds and use that knowledge to decode three-letter syllables that alone are nonsense words.

- Oral Reading Fluency (Mid-Grade 1 through Grade 3) – This subtest includes benchmark passages at each grade level that are used to measure accuracy and speed in reading graded passages.

- Oral Retelling Fluency (Mid-Grade 1 through Grade 3) – This optional assessment asks the student to tell as much as they can about a passage that they are asked to read.

- Word Use Fluency (Pre-Kindergarten through Grade 3) – This optional subtest is designed to assess vocabulary knowledge and expressive language for students in each grade level.

[4] Information is obtained from University of Oregon Center on Teaching and Learning online resource at http://dibels.uoregon.edu/dibelsinfo.php.

To streamline the assessment process for the purposes of the study, only the first five subtests were administered. The subtests were administered in an identical manner at both pre-test and post-test. To ensure the identifying and coding of reading variables occurred in an accurate manner for the purposes of the study, the on-site researchers used to conduct the DIBELS assessments received formal training and user's manuals from CREP researchers. On-site researchers conducted DIBELS with both treatment and control students at the beginning and end of LLI.

Intervention Fidelity

Leveled Literacy Intervention Observation Tool (LLIOT)

The LLIOT, developed by CREP researchers for the purposes of the study, involves a targeted, 30 minute observation of 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 20 items, the LLIOT is comprised of 3 subscales: Quality of LLI Implementation, which is designed to measure LLI teachers' implementation of the 10 main LLI lesson components; Literacy Instructional Strategies, which is designed to assess LLI teachers' use of general teaching strategies that should be present in a successful literacy intervention; and Learning Environment, which is designed to assess the quality of lesson factors such as organization, pacing, and the availability of materials. On-site researchers trained by CREP conducted observations of two intervention sessions with each participating LLI group, one near the beginning of the study period and one near the end, using the LLIOT. This observation data contributed to the evaluation of fidelity to the LLI model. 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 instruction on the instrument in a group session and participated in practice exercises.

LLI Data Management System Intervention Record

The Intervention Record in the LLI data management system was used for tracking student and teacher attendance, reasons for absence, student reading selections, and achievement level. This data management tool allows for individual or group reports to be created based on various criteria. CREP utilized the new online version of this data management program to access these intervention records, which provided an additional source of measurement of the fidelity of LLI implementation at each school.

School Support for Literacy: Instructional Staff Surveys

LLI Teacher Questionnaire (LLITQ)

An existing LLI teacher questionnaire that CREP had developed for a previous evaluation of LLI was modified and 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 and enthusiasm for literacy. The LLITQ consists of 21 items on a 5-point scale ranging from 1 (Strongly Disagree) to 5 (Strongly Agree), 5 items on a 4-point scale ranging from 0 (Not At All) to 3 (Extensively), 5 items on a 5-point scale ranging from 0 (Never) to 4 (Always), and 3 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 was administered to participating LLI teachers at the end of LLI.

Classroom Teacher Literacy Instruction Questionnaire (CTLIQ)

CREP also modified a previously developed teacher survey regarding literacy programs as a measure of 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 classroom teachers' self-reported literacy instructional practices and their perceptions of the core literacy program at their schools. The CTLIQ consists of 20 items on a 5-point scale ranging from 1 (Strongly Disagree) to 5 (Strongly Agree), 5 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 (Never) to 4 (Regularly), and 3 open-ended items regarding program strengths and areas for improvement as well as reasons to continue or not continue the school's current literacy program. The CTLIQ was administered to K-2 classroom teachers with either treatment or control students in the study at the end of LLI.

Focus Groups

Structured focus groups were conducted with both LLI teachers and on-site researchers at the end of the study period. Participating LLI teachers discussed their general view of LLI, logistical issues they encountered with implementing the system throughout the school year, LLI's strengths and areas for improvement, and their perceptions of the LLI online data management system, which was piloted by the publisher during the current study. On-site researchers also discussed their perceptions of LLI's strengths and areas for improvement as well as their general opinion of LLI, based on their random observations of LLI lessons.

Table 5 summarizes each of the research questions and the participants and provides the data sources and methodology used to investigate each question.

Table 5: Summary of Data Sources and Participants by Research Question

Research Questions	Participants	Data Sources	Method
1) What progress in literacy do students who receive LLI make compared to students who receive only regular classroom literacy instruction?	• LLI treatment and control students • LLI and classroom teachers	• Fountas & Pinnell Benchmarks • DIBELS • LLITQ • CTLIQ • LLI teacher focus groups	• Quantitative assessments of student progress in reading achievement • Qualitative assessment of student progress through teacher feedback
2) Was LLI implemented with fidelity to the developers' model?	• LLI teachers • On-site researchers	• LLIOT • LLI Data Management System • LLITQ • LLI teacher focus groups • On-site researcher focus groups	• Quantitative and qualitative assessments of LLI instructional strategies and delivery
3) What were LLI teachers' perceptions of LLI and its impact on their students' literacy?	• LLI teachers	• LLITQ • LLI teacher focus groups	• Quantitative and qualitative assessment of LLI teachers' perceptions regarding LLI's impact on their instruction and their students' literacy

Procedure

The current study extended from March 2009 through June 2010. In the spring of 2009, three CREP researchers were responsible for ensuring that the districts understood and agreed to participate in the study while implementing LLI as intended by the developers. A series of meetings was held with

key district-level administrators along with a presentation to all teachers who would be part of the study. Additionally, an LLI school coordinator was identified from the team of LLI teachers at each participating school to coordinate data collection activities with CREP and help ensure smooth LLI implementation. By the end of June 2009, orientation to the project along with district agreement was finalized. On- site researchers were also identified from a pool of local-area retired teachers during summer 2009, and a day-long training was organized to prepare them for the fall 2009 initiation of the evaluation. This on-site training included a detailed orientation to the LLI curriculum and evaluation as well as familiarization, thorough training, and practice with the assessment and observation instruments that the on-site researchers would administer (i.e., the DIBELS and LLIOT).

Prior to the first day of the 2009-2010 school year in each district, CREP researchers met separately with the teachers and the on-site researchers to finalize the timeline and logistics for pre-testing the first and second graders (kindergarteners received LLI during spring 2010). After the school year began, the schools in each district provided CREP with a list of first and second grade students that they had identified as eligible for LLI using their own selection criteria and whose parents had provided consent to participate in the study. Pre-testing of these students with the LLI Benchmarks and DIBELS began during the first three weeks of school. Subsequently, CREP conducted the randomization of the matched pairs of first and second graders based on demographic characteristics (i.e., gender, ethnicity, ELL status, special education status, and free/reduced lunch status) and pre-test LLI benchmark scores of instructional reading level. Students in the treatment group were then placed in LLI groups by LLI teachers, and the planned 90 days of LLI instruction for first and second graders began. Control group students did not receive LLI until the first and second grade evaluation period ended, and neither treatment nor control students received any additional pull-out literacy interventions during the study period. On- site researchers used the LLIOT to conduct two random observations of each first and second grade LLI group between October 2009 and February 2010, with one observation for each group occurring towards the beginning of LLI and one occurring towards the end. Post-tests with the LLI Benchmarks and DIBELS for the first and second grade students were completed in February for TCS and March for ECSDM. LLI school coordinators were also asked to encourage all first and second grade LLI teachers and first and second grade classroom teachers with students in the study to complete an online survey regarding either LLI or the regular classroom literacy program as applicable. CREP assisted in the online survey process by providing instructions and log-in information to all participating teachers. The first and second grade teacher surveys were administered in February and March 2010.

CREP researchers returned to both districts during February and March 2010 to conduct a refresher training on the DIBELS and LLIOT for the on-site researchers prior to the start of pre-testing the kindergarteners in the study. Follow-up visits with LLI teachers were also conducted at this time to discuss and address any concerns about the study thus far and to ascertain any issues related to LLI implementation and/or the online data management system that all LLI teachers were asked to use as part of the study. In late winter 2010, participating schools identified kindergarteners who were eligible for LLI and whose parents had provided consent for them to participate in the study. Pre-testing of these kindergarten students using the LLI Benchmarks and DIBELS began in February in TCS and March in ECSDM. Subsequently, these students were randomly assigned to treatment or control groups using the same randomization procedure that was utilized for first and second grade. Kindergartners in the treatment group received LLI beginning in February in TCS and April in ECSDM. As with the first and second grade groups, two random LLIOT observations were conducted for each kindergarten LLI group between March and May 2010. Post-testing of kindergarteners on the LLI Benchmarks and DIBELS was conducted during May 2010 in TCS and June 2010 in ECSDM.

During May and June 2010, end-of-year meetings were held with on-site researchers and LLI teachers 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, the online data management system, and participants' individual and collective views of LLI. Finally, LLI school coordinators were also asked to encourage all LLI and classroom teachers of kindergarten students in the study to complete an online survey regarding either LLI or the regular classroom literacy program as applicable. CREP assisted in the online survey process by providing instructions and log-in information to all participating teachers. The kindergarten teacher surveys were administered in May and June 2010. Table 6 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 6: Data Collection Summary

Type of Measure	Instrument	Timeline	Number Collected	Description
Student Achievement Measures	• LLI Benchmarks • DIBELS	August–October 2009 (1st & 2nd grade) February/March 2010 (K-2) May/June 2010 (K)	• 130 1st and 151 2nd grade pre/post-test LLI Benchmarks • 130 1st and 151 2nd grade pre/post-test DIBELS • 146 K pre/post-test LLI Benchmarks • 146 K pre/post-test DIBELS	• LLI benchmark and DIBELS testing for 1st and 2nd graders in both treatment and control groups was conducted as a pre-test in fall 2009 and as a post-test in winter 2010. • These same assessments were administered for kindergartners in both treatment and control groups as a pre-test in winter 2010 and as a post-test in spring 2010.
Surveys	• LLITQ • CTLIQ	February/March 2010 (1st & 2nd grade) May/June 2010 (K)	• 44 LLITQ's • 89 CTLIQ's	• Surveys were completed at the end of the 1st and 2nd grade LLI sessions by both LLI and classroom teachers, and again at the end of the kindergarten LLI sessions in the spring.
Observations	• LLIOT	October/Nov 2009 (1st & 2nd grade) January/February 2010 (1st & 2nd grade) March/April 2010 (K) April/May 2010 (K)	• 110 1st and 2nd grade LLIOT's • 50 K LLIOT's	• Trained on-site researchers observed all 1st and 2nd grade LLI groups twice in fall 2009/winter 2010. • These same researchers also observed all K LLI groups twice in spring 2010. Each observation lasted 30-45 minutes.
Focus Groups	• LLI Teacher Structured Focus Group • On-site Researcher Structured Focus Group	May/June 2010	• 2 LLI Teacher Focus Groups (1 per district) • 2 On-site Researcher Focus Groups (1 per district)	• LLI teacher focus groups were held in each district at the end of the school year to obtain qualitative feedback about LLI and students' progress from the LLI teachers. Each focus group lasted approximately 1 hour. • On-site researcher focus groups were held in each district at the end of the school year to obtain qualitative feedback about their observational experiences and measures used. Each focus group lasted approximately 1 hour.

Number of Days of LLI Instruction

Overall across both districts, the first and second grade LLI students received, on average, 72.9 days of instruction between August 2009 and March 2010, with a range of 40-90 days of instruction. Between February and June 2010, the kindergarten LLI students received, on average, 37.5 days of instruction, with individual students ranging in their attendance from 27 to 46 days.

In the five participating schools in Tift County Schools, the first and second grade LLI students received, on average, 78.9 days of instruction between August 2009 and February 2010, with individual students ranging in their attendance from 70 to 90 days. Between February and May 2010, the kindergarten LLI students in TCS received, on average, 36.7 days of instruction, with individual students ranging in their attendance from 27 to 46 days.

In the four participating schools in the Enlarged City School District of Middletown, the first and second grade LLI students received, on average, 68.1 days of instruction between October 2009 and March 2010, with individual students ranging in their attendance from 40 to 78 days. Between April and June 2010, the kindergarten LLI students in ECSDM received, on average, 38.7 days of instruction, with individual students ranging in their attendance from 36 to 42 days.

Results

The following section presents the results of the evaluation, discussed in relation to each instrument and each grade level. 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.

Preliminary Analyses

As the LLI benchmarks were scored in terms of alphabetic levels (i.e., pre-A, A, B, C, etc.), these outcomes first had to be recoded into numeric equivalents before analysis. Additionally, because some students were unable to reach the initial benchmark Level A as measured in the LLI benchmark system, we created a new category, pre-A benchmark level, in order to assign scores to those who were below Level A so those students could be included in the study. All benchmark outcomes were assigned numeric equivalents for each grade level before a series of mixed (i.e., "one between groups"/"one within groups") analysis of variance (ANOVA) procedures was conducted on the transformed measures to determine whether larger gains were observed for one of the two conditions overall (i.e., LLI/treatment vs. delayed-LLI/control) and for several demographic subgroups nested within the two conditions (e.g., ethnicity, special education status, English Language Learner status). Also, variations in the sample sizes across each analysis were seen due to limited cases of missing data. In the total sample, any cases with missing data could not be included in the analysis. Missing data resulted from several situations: 1) only cases with both pre-test and post-test data were able to be included in the analyses; 2) both achievement measures had "frustration" level cut-offs, which meant some students may not have had a score if they could not meet the minimum frustration level; and 3) students were allowed to voluntarily participate in the testing. Tests for normality of data and statistical assumptions (i.e., normal distribution; independence of measures) as well as measures of central tendency (i.e., means, standard deviations) were conducted on all outcomes for each grade level prior to the series of mixed ANOVAs.

Descriptive Student Achievement Results: Kindergarten LLI Benchmarks and DIBELS

Kindergarten LLI Benchmarks

On average after 38 days of LLI instruction, kindergartners who received LLI achieved a mean gain of 1.56 benchmark levels as compared to 0.78 benchmark levels for kindergartners who did not receive LLI. Also, kindergartners in LLI started, on average, below grade level in benchmark testing (i.e., pre-A = 0) but finished at a level between A and B, whereas their counterparts in the control group started near pre-A and finished around Level A. Thus, kindergartners in LLI finished the school year close to grade level in literacy (i.e., end-of-year K grade level goal = Level C). Also of note, English Language Learner (ELL), African American, and Hispanic students in LLI exceeded those in the control group. ELL students in LLI achieved a mean gain of about 1 ½ benchmark levels (M = 1.55) compared to a ½ benchmark level (M = 0.50) for ELL students not in LLI. African American LLI students also gained about 1 ½ benchmark levels (M = 1.44) while those in the control group only gained less than a benchmark level (M = 0.79). Finally, Hispanic students in LLI made the most gains—almost 2 benchmark levels (M = 1.76)—versus their counterparts in the control group who gained less than a benchmark level (0.70). Also, all three subgroups finished closer to grade level (i.e., Level C) than their counterparts who finished around Level A or below.

Kindergarten DIBELS

Overall, fewer significant gains were seen with the DIBELS outcomes. However, kindergartners in LLI significantly exceeded those who were not in LLI on nonsense word fluency (NWF) (M = 10.64% and M = 6.88%, respectively). Also, for phoneme segmentation fluency (PSF), ELL students in the treatment group (M = 46.72%) outperformed ELL students in the control group (M = 23.96%), *as well as* non-ELL students in both the treatment and control groups (M = 23.24% and 24.24%, respectively). Thus, kindergartners who participated in LLI showed more significant gains on subtests of the DIBELS as compared to those who did not have LLI.

Pre-test to Post-test Student Achievement Results: Mixed ANOVA Outcomes for Kindergarten LLI Benchmarks and DIBELS

Kindergarten LLI Benchmarks

As shown in Table 7, when the gains made by treatment and control group students were compared, highly significant differences favoring the treatment group were observed for the analysis involving all students ($F(1, 144) = 23.74$, $p < .001$, $\eta^2 = 0.14$). Although no statistically significant difference between conditions was observed when the gains made by White students ($F(1, 39) = 2.20$, *ns*, $\eta^2 = 0.05$) and students with a special education (SPED) designation ($F(1, 12) = 1.71$, *ns*, $\eta^2 = 0.13$) were analyzed, results favoring the treatment group were systematically observed when the analyses focused on African American students only ($F(1, 51) = 6.69$, $p < .05$, $\eta^2 = 0.12$), ELL students only ($F(1, 21) = 6.68$, $p < .05$, $\eta^2 = 0.24$), and Hispanic students only ($F(1, 48) = 16.22$, $p < .001$, $\eta^2 = 0.25$). As reflected in the magnitude of the effect sizes, the impact of treatment appeared to be especially robust with respect to literacy development in the latter two groups.

Table 7: Summary of Mixed ANOVA Results for Kindergarten LLI Benchmarks

Group/ Subgroup	Control Condition					Treatment Condition					F		η^2
		LLI Benchmark Pretest		LLI Benchmark Posttest			LLI Benchmark Pretest		LLI Benchmark Posttest				
	n	*M*	*SD*	*M*	*SD*	*n*	*M*	*SD*	*M*	*SD*			
Aggregate	70	0.26	0.53	1.04	1.00	76	0.20	0.46	1.76	0.89	23.74	***	0.14
SPED	4	0.00	0.00	0.75	0.96	10	0.30	0.67	1.80	0.79	1.71		0.13
ELL	12	0.25	0.45	0.75	0.97	11	0.27	0.47	1.82	1.25	6.68	*	0.24
African American	24	0.29	0.55	1.08	0.83	29	0.28	0.59	1.72	0.75	6.69	*	0.12
Hispanic/Latino	24	0.13	0.34	0.83	1.05	26	0.12	0.33	1.88	0.91	16.22	***	0.25
White/ Not Hispanic	21	0.38	0.67	1.29	1.10	20	0.20	0.41	1.60	1.05	2.20		0.05

***$p < .001$. **$p < .01$. *$p < .05$.

Kindergarten DIBELS

To contrast the pre- and post-test scores of kindergartners in treatment and control groups on the LLI Benchmarks, a second series of mixed Analysis of Variance (ANOVAs) was conducted on the means of four DIBELS measures of reading fluency. On three of four such measures—specifically, DIBELS Initial Sound Fluency (ISF, as shown in Table 9), Letter Naming Fluency (LNF, as shown in Table 10), and Phoneme Segmentation Fluency (PSF, as shown in Table 11)—no statistically significant differences in average student performance were observed either for groups of kindergarten students in the aggregate or for subgroups of kindergarten students disaggregated by ethnicity (African American, Hispanic, and White), ELL status, or special education status. However, as seen in Table 8 on the DIBELS measure of Nonsense Word Fluency (NWF), statistically significant differences were observed favoring treatment students. As shown in Table 8, such differences were observed for treatment students in the aggregate ($F(1, 139) = 5.97$, $p < .05$, $\eta^2 = 0.04$), as well for treatment students who were classified as ELL ($F(1, 21) = 4.90$, $p < .05$, $\eta^2 = 0.19$). Although the difference in performance on the NWF outcome for students with a special education designation (SPED) was not statistically significant, the effect size associated with the pre-test to post-test outcome suggests that the advantage of the treatment for these students over their counterparts in the control group was, nevertheless, a considerable one ($F(1, 12) = 1.55$, *ns*, $\eta^2 = 0.11$).

Table 8: Kindergarten DIBELS Nonsense Word Fluency Scores: % Correct

Group/ Subgroup	Control Condition					Treatment Condition					F		η^2
		NWF Pretest % correct		NWF Posttest % correct			NWF Pretest % correct		NWF Posttest % correct				
	n	*M*	*SD*	*M*	*SD*	*n*	*M*	*SD*	*M*	*SD*			
Aggregate	70	3.33	4.16	6.88	6.54	71	4.24	4.89	10.64	8.30	5.97	*	0.04
SPED	4	3.47	4.43	2.60	2.68	10	5.42	5.39	10.35	8.34	1.55		0.11
ELL	12	2.43	2.94	8.91	7.58	11	2.97	3.36	15.21	7.51	4.90	*	0.19
African American	24	3.41	4.06	6.89	5.69	27	3.78	4.74	10.47	7.75	3.66		0.07
Hispanic/Latino	24	2.69	3.13	6.39	7.04	24	4.37	4.48	11.60	8.46	2.17		0.05
White/ Not Hispanic	21	4.13	5.26	7.51	7.22	19	4.13	5.18	9.25	9.10	0.68		0.02

***$p < .001$. **$p < .01$. *$p < .05$.

Table 9: Kindergarten DIBELS Initial Sound Fluency Scores: % Correct

Group/ Subgroup	Control Condition					Treatment Condition						
		ISF Pretest % correct		ISF Posttest % correct			ISF Pretest % correct		ISF Posttest % correct			
	n	*M*	*SD*	*M*	*SD*	*n*	*M*	*SD*	*M*	*SD*	*F*	η^2
Aggregate	54	10.34	7.93	22.00	14.26	57	11.78	7.44	24.50	13.06	0.23	0.00
SPED	3	7.44	4.52	10.60	6.79	9	10.51	8.76	22.90	13.61	1.08	0.10
ELL	11	8.79	4.40	17.42	10.58	11	9.90	2.28	24.98	13.37	1.87	0.09
African American	21	10.29	7.03	21.36	15.64	24	9.81	7.08	22.21	13.76	0.13	0.00
Hispanic/Latino	15	9.25	4.46	22.40	14.95	17	12.60	5.05	28.32	12.55	0.31	0.01
White/ Not Hispanic	17	11.74	11.11	22.38	13.08	15	12.70	8.65	22.42	11.33	0.07	0.00

***p < .001. **p < .01. *p < .05.

Table 10: Kindergarten DIBELS Letter Naming Fluency Scores: % Correct

Group/ Subgroup	Control Condition					Treatment Condition						
		LNF Pretest % correct		LNF Posttest % correct			LNF Pretest % correct		LNF Posttest % correct			
	n	*M*	*SD*	*M*	*SD*	*n*	*M*	*SD*	*M*	*SD*	*F*	η^2
Aggregate	70	22.26	10.84	31.69	13.76	71	23.75	10.78	34.53	11.88	0.67	0.00
SPED	4	20.45	10.33	23.41	4.158	10	26.73	9.022	33.45	10.36	0.51	0.04
ELL	12	24.70	8.65	36.67	11.68	11	21.74	13.80	39.17	12.35	1.64	0.07
African American	24	22.20	8.83	30.87	14.74	27	23.20	11.88	33.30	12.92	0.25	0.01
Hispanic/Latino	24	21.78	10.78	33.94	14.02	24	23.86	11.73	36.74	11.17	0.06	0.00
White/ Not Hispanic	21	23.46	13.14	30.74	12.59	19	24.35	8.43	33.49	11.76	0.44	0.01

***p < .001. **p < .01. *p < .05.

Table 11: Kindergarten DIBELS Phoneme Segmentation Fluency Scores: % Correct

Group/ Subgroup	Control Condition					Treatment Condition						
		PSF Pretest % correct		PSF Posttest % correct			PSF Pretest % correct		PSF Posttest % correct			
	n	*M*	*SD*	*M*	*SD*	*n*	*M*	*SD*	*M*	*SD*	*F*	η^2
Aggregate	70	10.32	12.12	23.89	20.98	71	11.21	12.68	26.88	22.42	0.45	0.00
SPED	4	6.60	6.25	15.63	19.09	10	7.64	9.54	22.08	18.92	0.45	0.04
ELL	12	8.80	10.86	23.96	18.89	11	12.12	14.73	46.72	25.60	6.94 *	0.25
African American	24	7.93	6.72	21.70	19.58	27	7.66	10.06	17.64	17.55	0.79	0.02
Hispanic/Latino	24	12.56	13.65	26.85	23.12	24	14.18	13.23	38.89	24.26	3.33	0.07
White/ Not Hispanic	21	10.98	14.89	24.01	20.59	19	10.89	13.16	24.20	20.56	0.00	0.00

***p < .001. **p < .01. *p < .05.

Kindergarten Difference Score Analysis for LLI Benchmarks and DIBELS: Overall and by Subgroups

Additionally, we conducted analyses on treatment and control group difference scores (i.e., pre-test to post-test difference) in order to determine if any significant gain, or rate of change over time, was found for either group. From the pre- and post-test outcomes on the benchmark tests and DIBELS

measures, difference scores were computed and analyzed for treatment and control group students in the aggregate. Within this analysis, the performance of selected subgroups by students' special education and ELL statuses was contrasted as was the performance of student subgroups by ethnicity.

Overall , relative to the performance of control group students, significant differences were observed for the gains made by treatment group students on both the LLI benchmarks ($t(144) = 4.87$, $p < .001$, $d = .80$) and the DIBELS measure of Nonsense Word Fluency ($t(139) = 2.45$, $p < .05$, $d = .41$) (see Table 12).

Table 12: Overall Kindergarten Student Mean Difference Scores on LLI Benchmarks and DIBELS Subtests

Domain	Aggregate Control			Aggregate Treatment			t		d
	n	*M*	*SD*	*n*	*M*	*SD*			
Benchmarks	70	0.79	0.96	76	1.57	0.97	4.87	***	0.80
ISF	54	11.66	12.25	57	12.72	11.34	0.48		0.09
LNF	70	9.43	9.42	71	10.78	10.11	0.82		0.14
PSF	70	13.57	18.51	71	15.67	18.44	0.67		0.11
NWF	70	3.54	5.87	71	6.40	7.84	2.45	*	0.41

***$p < .001$. **$p < .01$. *$p < .05$.

With regard to subgroups, however, the only significant differences observed were the student gains on the DIBELS measure of Phoneme Segmentation Fluency. Specifically, significant differences were observed among students by experimental group and ELL status ($F(1, 136) = 6.53$, $p < .05$, $\eta^2 = 0.05$) as seen in Table 13, and among students by experimental group and ethnicity being African American or Hispanic ($F(1, 95) = 4.03$, $p < .05$, $\eta^2 = 0.04$) as shown in Table 14. With regard to the first result, follow-up testing indicated that the gains made by ELL students in the treatment group ($M = 34.60$, $SD = 20.78$) were superior to those made by non-ELL students in the treatment group ($M = 12.20$, $SD = 15.84$) and to those made by both ELL students ($M = 15.16$, $SD = 14.27$) and non-ELL students ($M = 13.43$, $SD = 19.49$) in the control group. As regards the latter outcome, the gains made by treatment group Hispanic students ($M = 24.71$, $SD = 21.02$) outpaced those made by African American students in the treatment group ($M = 9.98$, $SD = 12.38$).

Table 13: Kindergarten Student Mean Difference Scores on LLI Benchmarks and DIBELS Subtests: ELL Subgroup Comparison

Gain	Aggregate Control							Aggregate Treatment								
	Non ELL			ELL			d	Non ELL			ELL			F		d	η^2
	n	*M*	*SD*	*n*	*M*	*SD*		*n*	*M*	*SD*	*n*	*M*	*SD*				
Benchmarks	57	0.86	1.01	12	0.50	0.67	-0.38	65	1.57	0.93	11	1.55	1.21	0.58		-0.02	0.00
ISF	42	12.28	13.20	11	8.63	8.05	-0.3	46	12.16	10.87	11	15.09	13.45	1.35		0.26	0.01
LNF	57	8.93	9.84	12	11.97	7.40	0.33	60	9.56	9.19	11	17.44	12.64	1.21		0.82	0.01
PSF	57	13.43	19.49	12	15.16	14.27	0.09	60	12.20	15.84	11	34.60	20.78	6.53	*	1.36	0.05
NWF	57	2.90	5.72	12	6.48	6.18	0.63	60	5.32	7.66	11	12.25	6.31	1.19		0.94	0.01

***$p < .001$. **$p < .01$. *$p < .05$.

Table 14: Kindergarten Student Mean Difference Scores on LLI Benchmarks and DIBELS Subtests: Ethnicity Subgroup Comparison

Gain	Aggregate Control							Aggregate Treatment								
	African American			Hispanic			*d*	African American			Hispanic			*F*		*d*	η^2
	n	*M*	*SD*	*n*	*M*	*SD*		*n*	*M*	*SD*	*n*	*M*	*SD*				
Benchmarks	24	0.79	0.93	24	0.71	0.95	-0.09	29	1.45	0.91	26	1.77	0.91	1.22		0.36	0.01
ISF	21	11.07	12.36	15	13.15	14.21	0.16	24	12.40	12.55	17	15.72	11.82	0.04		0.28	0.00
LNF	24	8.67	11.92	24	12.16	7.84	0.35	27	10.10	8.13	24	12.88	12.18	0.03		0.28	0.00
PSF	24	13.77	17.90	24	14.29	18.44	0.03	27	9.98	12.38	24	24.71	21.02	4.03	*	0.88	0.04
NWF	24	3.47	5.47	24	3.70	6.96	0.04	27	6.69	6.41	24	7.23	9.45	0.01		0.07	0.00

***$p < .001$. **$p < .01$. *$p < .05$.

Descriptive Student Achievement Results: 1st Grade LLI Benchmarks and DIBELS

1st Grade LLI Benchmarks

On average after 73 days of LLI instruction, 1st graders who received LLI achieved a mean gain of 4.46 benchmark levels as compared to 2.63 benchmark levels for 1st graders who did not receive LLI. Also, 1st graders in LLI started generally below grade level in benchmark testing (i.e., Level A = 1) but finished at a level between E and F, whereas their counterparts in the control group started near Level A and finished around Level D. Thus, 1st graders in LLI finished their LLI sessions at the grade level mid-year goal in literacy (i.e., mid-year grade level goal for 1st grade = Levels E/F), while the control group students were still slightly behind. Also of note, African American and Hispanic students in LLI exceeded those in the control group. African American LLI students made the most gains—they gained about 5 ½ benchmark levels (M = 5.20) while those in the control group only gained about 2 ½ benchmark levels (M = 2.60). Finally, Hispanic students in LLI also made significant gains—about 4 benchmark levels (M = 4.18)—versus their counterparts in the control group who gained about 2 ½ benchmark levels (M = 2.57). Also, both subgroups finished at the grade level goal (i.e., Level E/F) compared to their counterparts in the control group who finished close to Level D. Of importance to note, the finding for African American 1st graders in LLI appears particularly robust and educationally significant. These LLI students finished the highest out of all subgroups as well as the aggregate—close to Level G—versus all others who finished between Levels C to F.

1st Grade DIBELS

Overall, similar significant differences between treatment and control groups were seen with the 1st grade DIBELS outcomes. 1st graders in LLI significantly exceeded those who were not in LLI on nonsense word fluency (NWF) (M = 22.00% and M = 17.00%, respectively). Also, for NWF, Hispanic students in the treatment group (M = 19.00%) outperformed their counterparts in the control group (M = 17.00%). Additionally, 1st graders who received LLI performed better than their counterparts on Oral Reading Fluency (M = 10.00% and M = 7.00%, respectively), as well as on Letter Naming Fluency (M = 17.00% and M = 11.00%, respectively). Thus, 1st graders who participated in LLI showed more significant gains on subtests of the DIBELS as compared to those who did not have LLI.

Pre-test to Post-test Student Achievement Results: Mixed ANOVA Outcomes for 1st Grade LLI Benchmarks and DIBELS

1st Grade LLI Benchmarks

As shown in Table 15, when the gains made by treatment and control group students on the LLI benchmarks were compared, significant differences favoring the treatment group were observed for the analysis involving all students ($F(1, 128) = 31.74$, $p < .001$, $\eta^2 = 0.20$) as well as for the analyses involving all subgroups except special education status ($F(1, 5) = 2.76$, *ns*) and ELL status ($F(1, 11) = 0.13$, *ns*).

Table 15: Summary of Mixed ANOVA Results for1st Grade LLI Benchmarks

Group/ Subgroup	Control Condition					Treatment Condition					*F*		η^2
		LLI Benchmark Pretest		LLI Benchmark Posttest			LLI Benchmark Pretest		LLI Benchmark Posttest				
	n	*M*	*SD*	*M*	*SD*	*n*	*M*	*SD*	*M*	*SD*			
Aggregate	65	1.32	1.03	3.95	2.37	65	1.37	1.18	5.83	2.27	31.74	***	0.20
SPED	3	1.33	0.58	2.67	0.58	4	1.00	1.41	4.25	3.30	2.76		0.36
ELL	10	1.40	0.97	5.00	2.21	3	1.33	0.58	5.33	1.53	0.13		0.01
African American	20	1.25	0.91	3.85	2.50	15	1.40	0.99	6.60	1.24	22.44	***	0.40
Hispanic/Latino	28	1.11	0.88	3.68	2.13	28	1.11	1.07	5.29	2.42	10.02	**	0.17
White/ Not Hispanic	17	1.76	1.30	4.53	2.62	20	1.60	1.43	6.00	2.66	5.90	*	0.14

***$p < .001$. **$p < .01$. *$p < .05$.

1st Grade DIBELS

To contrast the pre- and post-test scores of 1st graders in treatment and control groups on the LLI Benchmarks, a second series of mixed Analysis of Variance (ANOVAs) was conducted on the means of four DIBELS measures of reading fluency. On the DIBELS measure of Nonsense Word Fluency (NWF), statistically significant differences were observed favoring treatment students. As shown in Table 16, such differences were observed for treatment students in the aggregate ($F(1, 128) = 8.24$, $p < .01$, $\eta^2 = 0.06$), as well for Hispanic students in the treatment group ($F(1, 54) = 4.11$, $p < .05$, $\eta^2 = 0.07$).

Table 16: 1st Grade DIBELS Nonsense Word Fluency Scores: % Correct

Group/ Subgroup	Control Condition					Treatment Condition					*F*		η^2
		NWF Pretest % Correct		NWF Posttest % Correct			NWF Pretest % Correct		NWF Posttest % Correct				
	n	*M*	*SD*	*M*	*SD*	*n*	*M*	*SD*	*M*	*SD*			
Aggregate	65	0.10	0.07	0.17	0.09	65	0.11	0.07	0.22	0.11	8.24	**	0.06
SPED	3	0.08	0.05	0.26	0.11	4	0.11	0.05	0.16	0.09	4.93		0.52
ELL	10	0.09	0.06	0.21	0.07	3	0.07	0.07	0.17	0.10	0.14		0.01
African American	20	0.13	0.08	0.17	0.11	15	0.12	0.04	0.20	0.08	1.83		0.06
Hispanic/Latino	28	0.09	0.06	0.17	0.09	28	0.07	0.05	0.19	0.08	4.11	*	0.07
White/ Not Hispanic	17	0.10	0.07	0.19	0.09	20	0.13	0.09	0.28	0.14	2.16		0.06

***$p < .001$. **$p < .01$. *$p < .05$.

On the DIBELS measure of Oral Reading Fluency (ORF), statistically significant differences were observed favoring treatment students. As shown in Table 17, such differences were observed for treatment students in the aggregate ($F(1, 128) = 4.85$, $p < .05$, $\eta^2 = 0.04$), as well for treatment students who were White ($F(1, 35) = 8.70$, $p < .01$, $\eta^2 = 0.20$).

Table 17: 1st Grade DIBELS Oral Reading Fluency Scores: % Correct

Group/ Subgroup	Control Condition					Treatment Condition							
		ORF Pretest % Correct		ORF Posttest % Correct			ORF Pretest % Correct		ORF Posttest % Correct				
	n	*M*	*SD*	*M*	*SD*	*n*	*M*	*SD*	*M*	*SD*	*F*		η^2
Aggregate	65	0.04	0.04	0.11	0.10	65	0.04	0.03	0.14	0.10	4.85	*	0.04
SPED	3	0.03	0.02	0.08	0.03	4	0.04	0.02	0.11	0.03	1.54		0.24
ELL	10	0.06	0.07	0.20	0.13	3	0.04	0.03	0.13	0.06	0.70		0.06
African American	20	0.04	0.02	0.12	0.10	15	0.05	0.03	0.13	0.05	0.00		0.00
Hispanic/Latino	28	0.03	0.04	0.11	0.11	28	0.03	0.03	0.12	0.10	0.38		0.01
White/ Not Hispanic	17	0.04	0.04	0.10	0.09	20	0.04	0.03	0.18	0.12	8.70	**	0.20

***$p < .001$. **$p < .01$. *$p < .05$.

On the DIBELS measure of Letter Naming Fluency (LNF), statistically significant differences were observed favoring treatment students. As shown in Table 18, such differences were observed for treatment students in the aggregate ($F(1, 128) = 4.14$, $p < .05$, $\eta^2 = 0.03$). There was one statistically significant difference in LNF outcomes by ethnicity, which was counter to our hypotheses. ELL students in the control group outperformed those in the treatment group ($F(1,115) = 7.78$, $p < .01$, $\eta^2 = 0.41$). Finally, there were no significant effects found for the aggregate or subgroups for Phoneme Segmentation Fluency (PSF).

Table 18: 1st Grade DIBELS Letter Naming Fluency Scores: % Correct

Group/ Subgroup	Control Condition					Treatment Condition							
		LNF Pretest % Correct		LNF Posttest % Correct			LNF Pretest % Correct		LNF Posttest % Correct				
	n	*M*	*SD*	*M*	*SD*	*n*	*M*	*SD*	*M*	*SD*	*F*		η^2
Aggregate	65	0.31	0.13	0.42	0.19	65	0.30	0.15	0.47	0.17	4.14	*	0.03
SPED	3	0.29	0.10	0.42	0.16	4	0.21	0.12	0.28	0.04	0.36		0.07
ELL	10	0.32	0.10	0.51	0.18	3	0.27	0.11	0.28	0.10	7.78	*	0.41
African American	20	0.37	0.12	0.44	0.20	15	0.34	0.16	0.45	0.18	0.53		0.02
Hispanic/Latino	28	0.28	0.12	0.40	0.19	28	0.27	0.15	0.41	0.15	0.42		0.01
White/ Not Hispanic	17	0.28	0.13	0.43	0.18	20	0.33	0.13	0.56	0.16	3.25		0.09

***$p < .001$. **$p < .01$. *$p < .05$.

1st Grade Difference Score Analysis for LLI Benchmarks and DIBELS: Overall and by Subgroups

Additionally, we conducted analyses on treatment and control group difference scores (i.e, pre-test to post-test difference) in order to determine if any significant gain, or rate of change over time, was found for either group. From the pre- and post-test outcomes on the benchmark tests and DIBELS measures, difference scores were computed and analyzed for treatment and control group students in the aggregate. Within this analysis, the performance of selected subgroups by students' special education status and ELL status was contrasted as was the performance of student subgroups by ethnicity.

Overall, relative to the performance of control group students, statistically significant gains were made by treatment group students on both the LLI Benchmarks ($F = 31.97$, $p < .001$, $d = 1.26$) and the DIBELS measure of Nonsense Word Fluency ($F = 10.54$, $p = .001$, $d = .58$) (see Table 19).

Table 19: Overall 1st Grade Student Mean Difference Scores on LLI Benchmarks and DIBELS Subtests

Gain	Aggregate Control			Aggregate Treatment			*F*	*p*	*d*
	n	*M*	*SD*	*n*	*M*	*SD*			
Benchmarks	65	2.63	1.00	63	4.49	1.87	31.97	.000*	1.26
LNF	63	0.12	0.12	63	0.17	0.13	3.53	0.06	0.34
PSF	63	0.17	0.16	63	0.19	0.15	0.94	0.33	0.17
NWF	63	0.07	0.07	63	0.12	0.10	10.54	0.001*	0.58
ORF	63	0.08	0.08	63	0.10	0.08	3.47	0.07	0.33

***$p < .001$. **$p < .01$. *$p < .05$.

With regard to subgroups, there were several significant results. While not central to our hypotheses in the study, the analyses by the three ethnic subgroups (see Table 20) revealed there was a statistically significant difference among ethnicity for control students who did not receive LLI on the DIBLES measure for Nonsense Word Fluency (F = 6.24, p = .003, (d = -1.32 W vs. AA, d = -.33 W vs. H, d = .85 AA vs. H)), with White and Hispanic control students outperforming African American control students. Although there was a statistically significant difference for ethnicity on treatment group gains for Letter Naming Fluency (F = 3.35, p = .040, (d = -.69 W vs. AA, d = -.69 W vs. H, d = .17 AA vs. H)), and Oral Reading Fluency (F = 3.82, p = .028, (d = -.80 W vs. AA, d = -.65 W vs. H, d = .12 AA vs. H)), follow-up testing did not reveal significant differences among the individual ethnicities' group means (see Table 21).

Table 20: 1st Grade Student Mean Difference Scores on LLI Benchmarks and DIBELS Subtests: Ethnicity Subgroup Comparison for Control Students

	Aggregate Control													
Gain	White			African-American			Hispanic			*F*	*p*	d^1	d^2	d^3
	n	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*					
Benchmarks	17	2.76	2.05	20	2.60	1.93	28	2.57	1.75	0.06	0.94	-0.09	-0.11	-0.02
LNF	17	0.16	0.12	19	0.08	0.11	27	0.14	0.11	2.83	0.07	-0.76	-0.24	0.56
PSF	17	0.18	0.15	19	0.14	0.18	27	0.17	0.15	0.25	0.78	-0.20	-0.02	0.18
NWF	17	0.10	0.08	19	0.03	0.03	27	0.08	0.08	6.24	0.003*[a]	-1.32	-0.33	0.85
ORF	17	0.07	0.08	19	0.07	0.07	27	0.08	0.08	0.07	0.93	0.03	0.11	0.08

* $p < 0.05$

[a] White and Hispanic significantly higher than African American

[1] White vs. African-American; [2] White vs. Hispanic; [3] African-American vs. Hispanic

Table 21: 1st Grade Student Mean Difference Scores on LLI Benchmarks and DIBELS Subtests: Ethnicity Subgroup Comparison for Treatment Students

	Aggregate Treatment													
Gain	White			African-American			Hispanic			*F*	*p*	d^1	d^2	d^3
	n	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*					
Benchmarks	20	4.40	2.04	15	5.20	1.01	28	4.18	2.04	1.53	0.23	0.49	-0.11	-0.60
LNF	20	0.23	0.13	15	0.13	0.17	28	0.15	0.09	3.38	0.04*[b]	-0.69	-0.75	0.17
PSF	20	0.22	0.15	15	0.19	0.14	28	0.17	0.14	0.74	0.48	-0.23	-0.36	-0.14
NWF	20	0.15	0.14	15	0.09	0.10	28	0.11	0.06	1.31	0.28	-0.46	-0.32	0.30
ORF	20	0.14	0.10	15	0.08	0.03	28	0.09	0.08	3.82	0.03*[b]	-0.80	-0.65	0.12

* $p < 0.05$

[b] No significant post hoc tests.

[1] White vs. African-American; [2] White vs. Hispanic; [3] African-American vs. Hispanic

Descriptive Student Achievement Results: 2nd Grade LLI Benchmarks and DIBELS

2nd Grade LLI Benchmarks

On average after 73 days of LLI instruction, 2nd graders who received LLI achieved a mean gain of 4.64 benchmark levels as compared to 2.99 benchmark levels for 2nd graders who did not receive LLI. Also, 2nd graders in LLI started, on average, below grade level in benchmark testing (i.e., Level E = 5) but finished at Level J, whereas their counterparts in the control group started closer to Level F but only finished around Level I. Thus, 2nd graders in LLI finished the school year close to the grade level mid-year goal in literacy (i.e., mid-year grade level goal for 2nd grade = Level J/K). Also of note, while no significant effects were found for ELL students, a robust effect was found for students with a special education designation who received LLI. These students in the treatment group started around Level C and finished closer to Level H, while their counterparts in the control group started at Level D and finished around Level F. Also, regarding ethnicity subgroups, White students in LLI finished above their counterparts in the control group, gaining about 5 benchmark levels (M = 5.05) compared to about 3 benchmark levels (M = 3.14) in the control group. Additionally, African American and Hispanic students in LLI exceeded their counterparts in the control group. Of particular educational significance, African American LLI students finished at the highest level compared to all others—just above Level I; however, this was closely followed by the Hispanic LLI students who also finished slightly above Level I on average. The African American students in the treatment group gained about 4 ½ benchmark levels (M = 4.46) while those in the control group only gained about 2 ½ benchmark levels (M = 2.67). Finally, Hispanic students in LLI gained more than African American students in LLI (M = 4.53 and M = 4.46, respectively) while Hispanic students in the control group only gained about 3 benchmark levels.

2nd Grade DIBELS

Overall, no significant differences were found between treatment and control groups for 2nd grade on either DIBELS subtest that was administered as intended for 2nd graders (i.e., Nonsense Word Fluency and Oral Reading Fluency). While unexpected, this result may simply indicate that the 2nd grade DIBELS measures were not sufficiently aligned with the 2nd grade LLI curriculum or benchmarks to detect small effects, or changes, in DIBELS scores. However, it is also plausible that the lack of an overall effect may be due to district-level differences in these scores. One district appears to have made significant gains on the 2nd grade DIBELS tests compared to the other, but taken together, no overall effects were able to be seen (i.e., a wash-out effect from averaging across both districts' scores).

Pre-test to Post-test Student Achievement Results: Mixed ANOVA Outcomes for 2nd Grade LLI Benchmarks and DIBELS

2nd Grade LLI Benchmarks

As shown in Table 22, when the gains made by treatment and control group students on the LLI Benchmarks were compared, significant differences favoring the treatment group were observed for the analysis involving all students ($F(1, 149) = 22.58$, $p < .001$, $\eta^2 = 0.13$) as well as for the analyses involving all subgroups except ELL ($F(1, 19) = 0.80$, *ns*). As indicated by the magnitude of the effect sizes, the gains for students with a special education designation (SPED) were particularly strong ($\eta^2 = 0.47$) as well as for African American students ($\eta^2 = 0.17$).

Table 22: Summary of Mixed ANOVA Results for2nd Grade LLI Benchmarks

Group/ Subgroup	Control Condition					Treatment Condition							
		LLI Benchmark Pretest		LLI Benchmark Posttest			LLI Benchmark Pretest		LLI Benchmark Posttest				
	n	*M*	*SD*	*M*	*SD*	*n*	*M*	*SD*	*M*	*SD*	*F*		η^2
Aggregate	70	5.97	2.58	8.96	2.89	81	5.36	2.34	10.00	2.44	22.58	***	0.13
SPED	9	4.00	2.45	5.78	2.77	5	3.40	2.97	8.80	3.63	10.82	**	0.47
ELL	10	5.80	2.39	8.40	3.03	11	5.18	1.99	8.82	2.75	0.80		0.04
African American	24	6.33	2.62	9.00	3.43	30	5.67	2.12	10.13	2.56	10.46	**	0.17
Hispanic/Latino	22	5.41	2.48	8.64	2.63	30	5.50	2.54	10.03	2.65	4.38	*	0.08
White/ Not Hispanic	21	6.38	2.62	9.52	2.38	21	4.71	2.31	9.76	2.02	7.71	**	0.16

****p* < .001. ***p* < .01. **p* < .05.

2nd Grade DIBELS

To contrast the pre- and post-test scores of 2nd graders in treatment and control groups on the LLI Benchmarks, a second series of mixed Analysis of Variance (ANOVAs) was conducted on the means of two DIBELS measures of reading fluency. As mentioned earlier, neither the aggregate nor any of the subgroup DIBELS analyses for either Nonsense Word Fluency (Table 23) or Oral Reading Fluency (Table 24) were statistically significant.

Table 23: 2nd Grade DIBELS Nonsense Word Fluency Scores: % Correct

Group/ Subgroup	Control Condition					Treatment Condition						
		NWF Pretest % Correct		NWF Posttest % Correct			NWF Pretest % Correct		NWF Posttest % Correct			
	n	*M*	*SD*	*M*	*SD*	*n*	*M*	*SD*	*M*	*SD*	*F*	η^2
Aggregate	70	0.24	0.12	0.33	0.17	81	0.19	0.09	0.30	0.16	1.34	0.01
SPED	9	0.21	0.13	0.25	0.13	5	0.16	0.08	0.18	0.04	0.04	0.00
ELL	10	0.27	0.10	0.31	0.14	11	0.19	0.08	0.24	0.10	0.00	0.00
African American	24	0.26	0.15	0.34	0.21	30	0.16	0.08	0.27	0.14	0.43	0.01
Hispanic/Latino	22	0.25	0.09	0.34	0.13	30	0.20	0.09	0.32	0.19	0.71	0.01
White/ Not Hispanic	21	0.21	0.11	0.30	0.15	21	0.19	0.09	0.31	0.14	0.32	0.01

****p* < .001. ***p* < .01. **p* < .05.

Table 24: 2nd Grade DIBELS Oral Reading Fluency Scores: % Correct

Group/ Subgroup	Control Condition					Treatment Condition						
		ORF Pretest % Correct		ORF Posttest % Correct			ORF Pretest % Correct		ORF Posttest % Correct			
	n	*M*	*SD*	*M*	*SD*	*n*	*M*	*SD*	*M*	*SD*	*F*	η^2
Aggregate	70	0.13	0.08	0.21	0.11	81	0.11	0.07	0.21	0.09	1.28	0.01
SPED	9	0.09	0.05	0.14	0.06	5	0.09	0.07	0.17	0.11	0.73	0.06
ELL	10	0.12	0.07	0.22	0.11	11	0.09	0.03	0.18	0.08	0.18	0.01
African American	24	0.15	0.09	0.21	0.11	30	0.13	0.08	0.22	0.09	2.45	0.05
Hispanic/Latino	22	0.10	0.05	0.21	0.09	30	0.11	0.06	0.20	0.09	0.27	0.01
White/ Not Hispanic	21	0.13	0.09	0.23	0.11	21	0.11	0.06	0.22	0.08	0.84	0.02

****p* < .001. ***p* < .01. **p* < .05.

2nd Grade Difference Score Analysis for LLI Benchmarks and DIBELS: Overall and by Subgroups

Additionally, we conducted analyses on treatment and control group difference scores (i.e, pre-test to post-test difference) in order to determine if any significant gain, or rate of change over time, was found for either group. From the pre- and post-test outcomes on the benchmark tests and DIBELS measures, difference scores were computed and analyzed for treatment and control group students in the aggregate. Within this analysis, the performance of selected subgroups by students' special education and ELL statuses was contrasted as was the performance of student subgroups by ethnicity.

Overall, as compared to the control group, statistically significant differences were observed for the gains made by treatment group students on the LLI Benchmarks ($F = 22.58$, $p < .001$, $d = .78$) (see Table 25). However, no significant differences on either DIBELS subtest was found overall in 2nd grade, as shown in Table 25.

Table 25: Overall 2nd Grade Student Mean Difference Scores on LLI Benchmarks and DIBELS Subtests

	Aggregate Control			Aggregate Treatment			*F*	*p*	*d*
Gain	***n***	***M***	***SD***	***n***	***M***	***SD***			
Benchmarks	70	2.99	1.91	81	4.64	2.31	22.58	0.00*	0.78
NWF	70	0.09	0.14	81	0.11	0.13	1.33	0.25	0.19
ORF	70	0.09	0.06	81	0.10	0.07	1.27	0.26	0.19

***$p < .001$. **$p < .01$. *$p < .05$.

With regard to the subgroup of special education status shown in Table 26, the only statistically significant difference was observed on the LLI benchmarks, with the control students without a special education designation (non-SPED) scoring statistically significantly higher than the control students with a special education designation (SPED) ($F = 4.31$, $p = .04$, $d = -.75$). It should be noted that the sample size for the control students with a special education designation (N = 9) was very small. There were no statistically significant differences by special education status for students in the treatment group. However, again, the sample size for LLI students with a special education designation (N = 5) was very small and may not have been sufficient to detect effects. For the ELL subgroup (see Table 27), there were no statistically significant differences in gains between ELL and non-ELL students within either the treatment or control groups. Also, there were no statistically significant differences in gains between ethnicities within either the treatment or control groups (see Tables 28 and 29).

Table 26: 2nd Grade Student Mean Difference Scores on LLI Benchmarks and DIBELS Subtests: Special Education Subgroup Comparison

	Aggregate Control									Aggregate Treatment								
	Non SPED			SPED			*F*	*p*	*d*	Non SPED			SPED			*F*	*p*	*d*
Gain	***n***	***M***	***SD***	***n***	***M***	***SD***				***n***	***M***	***SD***	***n***	***M***	***SD***			
Benchmarks	61	3.16	1.94	9	1.78	1.20	4.31	0.04*	-0.75	76	4.59	2.28	5	5.40	2.97	0.57	0.45	0.35
NWF	61	0.10	0.14	9	0.04	0.16	1.18	0.28	-0.39	76	0.12	0.14	5	0.03	0.04	2.38	0.13	-0.72
ORF	61	0.09	0.07	9	0.06	0.03	2.18	0.14	-0.54	76	0.10	0.07	5	0.08	0.07	0.47	0.50	-0.32

* Significant at $p < 0.05$

Table 27: 2nd Grade Student Mean Difference Scores on LLI Benchmarks and DIBELS Subtests: ELL Subgroup Comparison

	Aggregate Control									Aggregate Treatment								
	Non ELL			ELL						Non ELL			ELL					
Gain	*n*	*M*	*SD*	*n*	*M*	*SD*	*F*	*p*	*d*	*n*	*M*	*SD*	*n*	*M*	*SD*	*F*	*p*	*d*
Benchmarks	60	3.05	1.85	10	2.60	2.37	0.47	0.50	-0.24	70	4.80	2.19	11	3.64	2.87	2.46	0.12	-0.51
NWF	60	0.10	0.14	10	0.05	0.15	1.14	0.29	-0.37	70	0.13	0.14	11	0.04	0.08	3.94	0.05	-0.65
ORF	60	0.08	0.07	10	0.10	0.04	0.57	0.45	0.26	70	0.10	0.07	11	0.09	0.07	0.22	0.64	-0.15

* Significant at $p < 0.05$

Table 28: 2nd Grade Student Mean Difference Scores on LLI Benchmarks and DIBELS Subtests: Ethnicity Subgroup Comparison for Control Students

	Aggregate Control													
	White			African American			Hispanic							
Gain	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*F*	*p*	d^1	d^2	d^3
Benchmarks	21	3.14	1.65	24	2.67	2.08	22	3.23	2.00	0.57	0.57	-0.26	0.05	0.28
NWF	21	0.09	0.13	24	0.09	0.15	22	0.08	0.15	0.03	0.97	-0.05	-0.07	-0.02
ORF	21	0.10	0.06	24	0.06	0.07	22	0.10	0.06	2.92	0.06	-0.54	0.13	0.66

* Significant at $p < 0.05$
[1] White vs. African-American
[2] White vs. Hispanic
[3] African-American vs. Hispanic

Table 29: 2nd Grade Student Mean Difference Scores on LLI Benchmarks and DIBELS Subtests: Ethnicity Subgroup Comparison for Treatment Students

	Aggregate Treatment													
	White			African American			Hispanic							
Gain	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*F*	*p*	d^1	d^2	d^3
Benchmarks	21	5.05	2.67	30	4.47	2.00	30	4.53	2.37	0.44	0.65	-0.26	-0.21	0.03
NWF	21	0.12	0.13	30	0.11	0.12	30	0.12	0.16	0.05	0.95	-0.06	0.03	0.08
ORF	21	0.11	0.06	30	0.09	0.07	30	0.09	0.06	0.64	0.53	-0.30	-0.30	0.02

* Significant at $p < 0.05$
[1] White vs. African-American
[2] White vs. Hispanic
[3] African-American vs. Hispanic

Implementation Fidelity: LLIOT

Descriptive Results

The Leveled Literacy Intervention Observation Tool (LLIOT) involved a targeted, 30-minute observation of LLI implementation and instructional strategies (n = 160 observations). Table 30 illustrates the frequencies for each item on the LLIOT, as observed during the visits. The results from the LLIOT revealed that 5 of the 10 LLI lesson components were rated "Acceptable" or "Excellent" over 90% of the time, indicating a high level of implementation fidelity across both districts. The highest rated lesson components (i.e., those demonstrating the highest degree of implementation fidelity) included phonics/word work, reading a new book, and rereading, which were rated "Acceptable" or "Excellent" 97.6%, 95.7%, and 95.0% of the time, respectively. The lowest rated lesson components (i.e., those demonstrating the lowest degree of implementation fidelity) included classroom and home connections, which were not observed 51.9% and 22.5% of the time, respectively. Teachers were also rated highly on their use of literacy instructional strategies, such as modeling and encouraging fluent oral reading (96.9% "Acceptable" or "Excellent") and appropriate reading strategies (95.7%) and assisting students in problem-solving (95.6%). Further, in the majority of observed lessons, instructional materials were readily available; the lesson was well-organized; and students were engaged and attentive (100.0%, 99.4%, and 98.1% "Acceptable" or "Excellent," respectively). Overall, observers perceived that the lesson was delivered as designed 96.3% of the time. All items can be found in Table 30 below.

Table 30: Leveled Literacy Intervention Observation Tool Response Frequencies (n = 160)

	Percent Responded			
Item	Excellent	Acceptable	Needs Improvement	Not Observed
Quality of LLI Implementation				
Rereading (shared or independent)	64.4	30.6	3.8	0.6
Assessment using Reading Record*	36.3	11.3	0.0	52.5
Phonics/word work (e.g., parts of words; single letter sounds; letter, word, or picture cards; magnetic letters)	66.3	31.3	1.3	0.6
Writing about reading (interactive, dictated, or independent)*	35.6	14.4	0.6	48.1
New book - Introducing new text	70.0	25.0	0.0	5.0
New book - Reading (shared or independent)	66.3	29.4	0.6	3.1
New book - Monitoring and supporting students as needed	71.3	21.3	1.3	5.6
New book - Discussing and revisiting the story	62.5	26.3	0.6	10.0
Classroom connection	32.5	13.8	0.0	51.9
Home connection	43.1	33.8	0.0	22.5
Literacy Instructional Strategies				
Teacher models, encourages, and provides opportunities for fluent oral reading.	68.8	28.1	1.3	1.9
Teacher introduces vocabulary words (e.g., high frequency, story-specific words).	56.3	35.0	3.1	5.6
Teacher emphasizes understanding/comprehension of what is read.	66.3	28.8	0.6	4.4
Teacher models and encourages students to use appropriate reading strategies (e.g., phonemic awareness).	66.9	28.8	0.0	4.4
Teacher engages students in conversation about the text.	70.6	23.1	0.6	5.0
Teacher assists students in problem-solving.	65.6	30.0	0.0	4.4

Table 30, continued

Item	Percent Responded			
	Excellent	Acceptable	Needs Improvement	Not Observed
Learning Environment				
Lesson is well organized.	80.0	19.4	0.6	0.0
Teacher appropriately paces lesson components.	68.1	29.4	2.5	0.0
Teacher engages in ongoing assessment of student learning (e.g., questioning, providing feedback/corrective instruction, checking responses).	69.4	30.0	0.0	0.6
Students are actively engaged.	75.6	22.5	1.3	0.0
Instructional modifications are observed when needed.	57.5	33.1	0.0	9.4
Instructional materials needed to implement lesson are readily available.	85.6	14.4	0.0	0.0
The lesson is delivered as designed.	61.3	35.0	3.1	0.0

Note: Item percentages may not total 100% because of missing input from some respondents.
*These items were each observed approximately 50% of the time because they are only implemented during even-numbered lessons.

The LLIOT also included items designed to describe the groups observed, which are summarized in Table 31. Results from these items indicated that the three grade levels were observed equally and that the majority of observed groups took place in a designated intervention area. Further, the majority of observed groups had three students and lasted approximately 30 minutes, which was consistent with LLI's design. Finally, an equal number of even- and odd-numbered lessons were observed. All items can be found in Table 31 below.

Table 31: Leveled Literacy Intervention Observation Tool Summary Items (n = 160)

Item	Percent Responded
Grade Level	
K	31.3
1	31.3
2	41.3
Location of Group	
Intervention Area	94.4
Classroom	5.0
Other	0.0
Number of Students in Group	
1	2.5
2	20.6
3	75.0
4 or more	1.9
Total Instructional Minutes	
Less than 25	3.1
25 – 35	62.5
More than 35	32.5
LLI Lesson Number	
Even	49.4
Odd	50.6

Note: Item percentages may not total 100% because of missing input from some respondents.

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. Of the 163 comments, 71.8% were related to the lesson resources and materials, 41.1% were related to student participation and engagement in the lesson, 38.7% were related to the quality of literacy instruction, and 20.9% were related to the success of the lesson[5]. In most of the comments, observers reported that a wide variety of instructional materials were readily available, the quality of instruction (e.g., pacing, teaching strategies, organization) was high; the students were actively engaged and enthusiastic; and the lesson was implemented successfully. However, multiple observers commented that the lesson could not adequately be completed within the designated 30-minute timeframe. Sample comments from the observers are provided below.

> "Lesson #81 was observed. All materials were readily available. Students were on task and eager to interact with all aspects of the lesson. The lesson went very well."
>
> "Lesson #14. Students very excited about their writing in their books. Enthused about their new book and related to what they saw on their way to school. Excited about what they were going to write for homework. It takes longer than 30 minutes to do a lesson."
>
> "While a student was being assessed the other two students were actively engaged in rereading books. They stayed focused and it was evident that they were quite pleased with themselves. Students seemed used to routine and complied with all directives. All materials needed were readily available for lesson 34."

Comparative Pre-test to Post-test Results

The LLIOT was conducted at both the beginning and the end of LLI for each of the observed groups in order to measure changes in implementation over time. For 1st and 2nd grade, pre-test observations were conducted in October and November 2009, and post-test observations were conducted in January and February 2010. For kindergarten, pre-test observations were conducted in March and April 2010, and post-test observations were conducted in April and May 2010. Results are summarized by grade level below.

Kindergarten

For each of the two occasions on which 25 kindergarten LLI groups were observed, the means and standard deviations presented in Table 32 were computed on the three subscales of the LLIOT. Subsequently, three independent *t*-tests that contrasted teacher behaviors at times one and two were conducted on the pairs of means obtained on the LLIOT's ten-item "Quality of LLI Implementation" scale ($t(23) = 0.81$, $p = 0.422$), its six-item "Literacy Instructional Strategies" scale ($t(23) = -0.40$, $p = 0.692$), and its seven-item "Learning Environment" scale ($t(23) = -0.18$, $p = 0.860$), with no statistically significant differences observed for any of the three comparisons. The average rating was between "Acceptable" (2.00) and "Excellent" (3.00) for each subscale at both time points.

1st Grade

The descriptive statistics and independent *t*-test results for each of the three LLIOT subscales for the 25 observed 1st grade groups are presented in Table 32. There were no significant differences between the pre-test and the post-test observations for two of the subscales: "Quality of LLI Implementation" ($t(48) = 0.05$, $p = 0.962$) and "Literacy Instructional Strategies" ($t(48) = 1.41$, $p = 0.165$).

[5] Percentages do not add up to 100% because a single comment may have addressed more than one theme.

However, scores on the "Learning Environment" scale did significantly improve from pre-test to post-test ($t(48) = 2.22$, $p < 0.05$). For all subscales, the average rating was between "Acceptable" (2.00) and "Excellent" (3.00) at both time points.

2nd Grade

Similar to 1st grade, the results of the three independent *t*-tests for the 33 observed 2nd grade groups revealed that there were no significant differences between the pre-test and the post-test observations for the subscales "Quality of LLI Implementation" ($t(55) = 1.35$, $p = 0.183$) and "Literacy Instructional Strategies" ($t(64) = 1.61$, $p = 0.113$). However, scores on the "Learning Environment" scale did significantly improve from pre-test to post-test ($t(49) = 2.47$, $p < 0.05$). The average rating was between "Acceptable" (2.00) and "Excellent" (3.00) for each subscale at both time points. Descriptive statistics and independent *t*-test results are summarized in Table 32 below.

Table 32: Independent T-Test Results for LLIOT Subscales by Grade Level

	Pre-Test		Post-Test				
LLIOT Subscale	***M***	***SD***	***M***	***SD***	***t***	***p***	***d***
Kindergarten (n = 25)							
Quality of LLI Implementation	2.26	0.77	2.11	0.46	0.81	0.422	-0.23
Literacy Instructional Strategies	2.79	0.32	2.83	0.27	-0.40	0.692	0.12
Learning Environment	2.75	0.34	2.77	0.34	-0.18	0.860	0.05
1st Grade (n = 25)							
Quality of LLI Implementation	2.02	0.61	2.03	0.56	0.05	0.962	0.02
Literacy Instructional Strategies	2.44	0.49	2.62	0.41	1.41	0.165	0.41
Learning Environment	2.62	0.33	2.81	0.27	2.22*	0.031	0.64
2nd Grade (n = 33)							
Quality of LLI Implementation	2.12	0.60	2.29	0.41	1.35	0.183	0.34
Literacy Instructional Strategies	2.29	0.84	2.57	0.58	1.61	0.113	0.39
Learning Environment	2.46	0.53	2.72	0.30	2.47*	0.017	0.61

*Statistically significant at $p < 0.05$

LLI Teacher Survey: LLITQ

The Leveled Literacy Intervention Teacher Questionnaire (LLITQ) was administered online to LLI teachers at the end of the study period as a general measure of their implementation and perceptions of LLI (n = 44 respondents). Table 33 illustrates the frequencies of responses for each item on the LLITQ. Most of the respondents reported positive perceptions of LLI and its implementation in their individual schools. LLI teachers indicated that they had a good understanding of LLI; received support in implementing LLI from their district, school administration, and other school staff; and perceived a positive impact on student achievement and student attitudes towards literacy. LLI teachers also reported a positive impact of LLI on their reading instruction, particularly their understanding of the role of comprehension and phonics/phonemic awareness in the reading process and the relationship of leveled texts to successful reading. Further, LLI teachers reported implementing the LLI system with a high degree of fidelity; the majority of teachers indicated that they met with their groups daily for at least 30 minutes, followed the lesson design, and implemented both reading and writing activities.

Overall, LLI teachers were most likely to agree that they understood the goals and implementation procedures for LLI (each 97.7% "Strongly Agree" or "Agree"), that LLI positively impacts student literacy achievement (97.7%), and that their districts and other teachers within their schools were supportive of LLI (each 93.2%). LLI teachers were least likely to agree that the parents of their LLI

students participated in home literacy activities with their children (45.4% "Disagree" or "Strongly Disagree"), that their schools had sufficient faculty and staff to provide LLI to all students who needed it (36.4% "Disagree" or "Strongly Disagree"), and that LLI helped their students with special needs[6] and ELL students (38.6% and 29.6% "Somewhat" or "Not At All," respectively). All of the surveyed teachers (100.0%) agreed that their school should continue LLI. All items can be found in Table 33 below.

Table 33: Leveled Literacy Intervention Teacher Questionnaire Response Frequencies (n = 44)

Item	Percent Responded		
	Strongly Agree/Agree	Neutral	Disagree/Strongly Disagree
I understand the goals of the Leveled Literacy Intervention system.	97.8	2.3	0.0
I have received adequate professional development for implementing LLI.	90.9	4.5	4.5
I have a thorough understanding of how to implement LLI.	97.7	0.0	2.3
Guidance and support is provided by our school staff to help us implement LLI.	86.3	9.1	4.5
LLI has positively impacted LLI student achievement.	97.7	2.3	0.0
LLI teachers are given sufficient planning time to implement the system.	70.5	11.4	18.1
Students who receive LLI in this school are more enthusiastic about reading, writing, and learning because of LLI.	75.0	20.5	4.5
Our school has sufficient faculty and staff to provide LLI to all students who need the intervention.	54.5	9.1	36.4
The administration protects the time needed for daily uninterrupted LLI teaching.	88.6	9.1	2.3
Parents participate in LLI home literacy activities with their child(ren).	22.7	31.8	45.4
Teachers in this school are generally supportive of LLI.	93.2	6.8	0.0
Ongoing communication exists between LLI teachers and classroom teachers.	93.2	2.3	4.6
LLI teachers are encouraged to communicate concerns, questions, and constructive ideas regarding the system to school staff or administration.	84.1	2.3	13.7
LLI allows for teachers to provide differentiated instruction to address the varying strengths and needs of students.	68.2	11.4	20.5
Instructional materials (books, assessments, and other resources) needed to implement LLI are readily available.	86.4	6.8	6.8
The faculty, staff, and administration in my school believe that all children can learn to read and write.	88.6	9.1	2.3
LLI is aligned with state and district reading and language arts standards.	93.2	4.5	2.3
LLI training has improved my reading instruction.	72.7	22.7	4.6
Our principal is an effective instructional leader.	93.1	2.3	4.6
LLI adequately prepares our students for state assessments.	59.1	38.6	2.3

[6] On both the LLI teacher survey and the classroom teacher survey, the term "students with special needs" was used to describe both students with a special education designation and students who had not yet been classified.

Table 33, continued

Item	Percent Responded		
	Strongly Agree/Agree	Neutral	Disagree/Strongly Disagree
Because of LLI, I have a greater understanding of…			
The reading process.	59.1	22.7	18.2
The characteristics of leveled books and their relationship to successful reading.	72.7	15.9	11.4
How to improve children's vocabulary and oral language skills.	70.5	13.6	15.9
The role of fluency in effective reading.	65.9	15.9	18.2
The role of phonics and phonemic awareness in the reading process.	72.7	15.9	11.4
The role of comprehension in successful reading.	84.1	9.1	6.8
How to improve children's writing strategies.	63.6	22.7	13.6

Item	Percent Responded		
	Extensively/ Sufficiently	Somewhat	Not at all
To what degree does your school administration support your efforts as an LLI teacher?	93.1	4.5	0.0
To what degree does the district support your efforts as an LLI teacher?	93.2	4.5	0.0
To what degree does your teaching schedule allow time to implement LLI effectively?	79.5	18.2	0.0
To what extent do you feel LLI has helped your English Language Learner students?	63.7	27.3	2.3
To what extent do you feel LLI has helped your students with special needs?	56.8	31.8	6.8

Item	Percent Responded		
	Always/ Frequently	Sometimes/ Rarely	Never
How often did your LLI group lessons last 30 minutes or more?	86.4	11.4	0.0
Were you able to meet every day with your LLI group(s)?	93.2	4.5	0.0
How often did you follow the LLI lessons exactly as instructed in the Lesson Guide?	75.0	20.5	0.0
How often were you able to implement LLI reading activities, such as phonics/word work and guided reading?	90.9	6.8	0.0
How often were you able to implement LLI writing activities, such as interactive writing?	90.9	6.8	0.0

Item	Percent Responded
Do you think your school should continue the Leveled Literacy Intervention system?	
Yes	100.0
No	0.0

Note: Item percentages may not total 100% because of missing input from some respondents.

The LLITQ also invited LLI teachers to share open-ended comments regarding the strengths and weaknesses of LLI and the reasons that their schools should continue or not continue using the LLI system. With respect to strengths, the majority of comments (81.0%)[7] focused on the LLI resources and materials. Teachers particularly liked the leveled texts, commenting that they were well-written and engaging to students. Several comments (35.7%) also focused on instructional strategies, including small group format and structure and consistency of lessons, and a few (4.8%) mentioned student achievement as a strength of LLI.

[7] Percentages do not add up to 100% because a single comment may have addressed more than one theme.

With regard to areas of improvement for LLI, teachers again mentioned the system's resources the majority of the time (80.5%), citing inconsistency among materials (e.g., differences between activities, books, and CD resources) and materials that did not appear to be level-appropriate. Teachers also frequently mentioned lesson design (34.1%), stating that the system was too fast-paced and that they could not cover all of the lesson material during the designated 30-minute timeframe.

When asked why their school should continue using the LLI system, teachers equally cited student achievement (41.5%) and benefit to ELL students and students with special needs (41.5%) – an interesting finding when compared to teachers' responses on the first part of the survey, in which they rated LLI as helping ELL students and students with special needs "Somewhat" or "Not At All" approximately one-third of the time (29.6% and 38.6%, respectively). No teachers provided a reason for their school to discontinue the use of LLI. Sample comments from the LLI teachers are provided below.

What are the strengths of the Leveled Literacy Intervention system?

> **"**I love the books that are provided for the program. They are current and fun. The children are excited about reading them. The children really enjoy doing the plays during familiar reading time."

> **"**It provides a safety net for those students who still need extra support in addition to the regular classroom instruction. Students enjoyed coming to LLI class and were motivated and excited about reading each day. It helped students to improve comprehension skills by being in a small group and learning to take risks in answering questions."

What areas of the Leveled Literacy Intervention system could be improved?

> "It is impossible to get writing and a running record done on the same day along with everything else in a lesson within the 30 minutes allotted. Thirty minutes is often not enough time to do a lesson that involves every task in each of the lesson components."

> "The levels move up very fast for some of the struggling students. It was hard to do everything in the length of time for the slower working student. Perhaps the lessons need to stretch into two days or stretch the levels into three weeks instead of two."

Why should your school continue or not continue the Leveled Literacy Intervention system?

> "[It] has all of the components needed to effectively improve student achievement."

> "I believe LLI meets struggling students' needs very effectively. All of the components work together to promote the reading process."

Classroom Teacher Survey: CTLIQ

The Classroom Teacher Literacy Instruction Questionnaire (CTLIQ) was administered online at the end of the study period to K-2 classroom teachers with students in the current study (either treatment or control) as a general measure of classroom teachers' literacy instructional strategies and perceptions of the core literacy program at their schools (n = 89 respondents). Table 34 illustrates the frequencies of responses for each item on the CTLIQ. Results from the CTLIQ revealed that classroom teachers of both treatment and control students were most likely to provide individual or small-group reading instruction, integrate vocabulary and comprehension into their literacy instruction, and utilize high-quality literature to read to students and engage them in interactive discussions about the text (each rated "Regularly" or "Frequently" 100.0% of the time). Teachers also frequently reported using guided reading instruction with leveled texts (95.5% "Regularly" or "Frequently") and writing activities (94.4%) and teaching phonological awareness to their students (93.2%). Teachers were least likely to report utilizing whole-class reading instruction (rated "Not At All" 7.9% of the time) and assigning home literacy activities for students to complete with parents (6.7%).

Overall, classroom teachers reported a positive perception of their school's literacy program. Teachers were most likely to agree that they understood the goals of their literacy program, that it was aligned with state and district reading/language arts standards, and that their faculty, staff, and administration believed that all students could learn to read and write (93.3%, 93.3%, and 92.1% "Strongly Agree" or "Agree," respectively). Similar to the LLI teachers, classroom teachers were least likely to agree that the parents of their students participated in home literacy activities with their children (27.0% "Disagree" or "Strongly Disagree"), that their schools had sufficient faculty and staff to fully implement their literacy program (25.9% "Disagree" or "Strongly Disagree"), and that their literacy program helped their students with special needs and ELL students (39.4% and 28.1% "Somewhat" or "Not At All," respectively). The majority of surveyed teachers (83.1%) agreed that their school should continue their current literacy program. All items can be found in Table 34 below.

Table 34: Classroom Teacher Literacy Instruction Questionnaire Response Frequencies (n = 89)

Item	Percent Responded		
	Strongly Agree/Agree	Neutral	Disagree/Strongly Disagree
I understand the goals of our school's literacy program.	93.3	4.5	2.2
I have received adequate professional development for implementing our school's literacy program.	77.6	7.9	13.5
I have a thorough understanding of how to implement our school's literacy program.	76.4	12.4	11.2
Guidance and support is provided by our school staff to help us implement our literacy program.	77.6	16.9	5.6
Our literacy program has positively impacted student achievement.	74.1	19.1	5.6
Teachers are given sufficient planning time to fully implement our school's literacy program.	55.1	20.2	24.7
Students in this school are more enthusiastic about reading, writing, and learning because of our literacy program.	67.4	24.7	7.8
Our school has sufficient faculty and staff to fully implement its literacy program.	65.2	9.0	25.9
The administration protects the time needed for daily uninterrupted literacy instruction.	85.4	10.1	4.5
Parents participate in home literacy activities with their child(ren).	32.5	39.3	27
Teachers in this school are generally supportive of our literacy program.	79.7	18.0	2.2

Table 34, continued

Item	Percent Responded		
	Strongly Agree/Agree	Neutral	Disagree/Strongly Disagree
Ongoing communication exists between LLI teachers and classroom teachers.	73.1	18.0	8.9
Teachers are encouraged to communicate concerns, questions, and constructive ideas regarding our literacy program to school staff or administration.	78.6	13.5	7.9
Our literacy program allows for teachers to provide differentiated instruction to address the varying strengths and needs of students.	83.2	11.2	5.6
Instructional materials (books, assessments, and other resources) needed to implement our literacy program are readily available.	74.1	13.5	12.4
The faculty, staff, and administration in my school believe that all children can learn to read and write.	92.1	6.7	1.1
Our literacy program is aligned with state and district reading and language arts standards/frameworks.	93.3	5.6	1.1
Professional development for our school's literacy program has improved my reading instruction.	76.4	15.7	7.8
Our principal is an effective instructional leader.	77.5	14.6	6.7
Our literacy program adequately prepares our students for state assessments.	71.9	20.2	7.8

Item	Percent Responded		
	Extensively/ Sufficiently	Somewhat	Not at all
To what degree does your school administration support your efforts to implement your school's literacy program?	87.6	12.4	0.0
To what degree does the district support your efforts to implement your school's literacy program?	83.1	16.9	0.0
To what degree does your teaching schedule allow time to implement your school's literacy program effectively?	75.3	22.5	2.2
To what extent do you feel your school's literacy program has helped your English Language Learner students?	69.6	23.6	4.5
To what extent do you feel your school's literacy program has helped your students with special needs?	59.5	31.5	7.9

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)
Students participate in whole group reading instruction.	71.9	20.2	7.9
Students participate in small group or individual reading instruction.	100.0	0.0	0.0
I provide guided reading instruction using leveled texts for groups of students with similar learning needs.	95.5	3.3	1.1
Students meet in small, heterogeneous groups to discuss the books that they are reading.	68.5	25.8	5.6
Students participate in writing activities, such as mini-lessons, independent writing, conferencing, and sharing.	94.4	4.5	0.0
I provide opportunities to develop oral reading fluency (e.g., shared reading, partner reading).	89.9	10.1	0.0
I teach phonological awareness (sound patterns, rhymes, etc.) to my students.	93.2	6.7	0.0
I integrate both vocabulary and comprehension into my literacy instruction and activities.	100.0	0.0	0.0

Table 34, 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)
I read high-quality children's literature (e.g., fiction, non-fiction, poetry) to my students and engage them in interactive discussions about the text.	100.0	0.0	0.0
I assign students home literacy activities to encourage parent participation.	51.6	39.3	6.7

Item	Percent Responded
Do you think your school should continue with the current literacy program?	
Yes	83.1
No	9.0

Note: Item percentages may not total 100% because of missing input from some respondents.

The CTLIQ also invited classroom teachers to share open-ended comments regarding the strengths and weaknesses of their school's literacy program and the reasons that their school should continue or not continue the literacy program. Regarding program strengths, most comments (36.8%)[8] cited instructional strategies, particularly the opportunities to provide differentiated and small group instruction. One-third of the comments (33.3%) focused on the resources and materials available to both students and teachers (e.g., a variety of books), while several comments (22.8%) stated that the program is individualized and allows students to succeed at their own level. A few comments (3.5%) also mentioned school climate (e.g., teachers/staff, support from literacy coaches and administration).

With respect to areas for improvement, teachers most frequently mentioned materials (50.0% of comments), particularly the need for more guided reading texts. The comments also recommended increased help for at-risk students (11.5%), more professional development (9.6%), improved instructional strategies (e.g., guided reading, one-on-one time, phonics instruction; 9.6%), and additional resources (e.g., better coaching/support, more intervention staff, improved time management; 7.7%). A few comments (5.8%) also mentioned LLI, with one teacher stating that all teachers should receive training in LLI and two teachers requesting better communication between LLI teachers and classroom teachers.

When asked why their school should continue its current literacy program, teachers most frequently cited a positive impact on student achievement (41.4% of comments), while some teachers voted not to continue the current literacy program because it does not meet the needs of all students (5.1% of comments). In 8.6% of comments, teachers voted specifically to continue the use of LLI, with one classroom teacher stating, "…I believe that children that struggle would give up hope in the realm of reading without the LLI program." Some teachers stated that their school must find a way to serve more students with LLI if it is to continue (5.2% of comments), while others reported that they did not know enough about LLI to offer an opinion about whether their school should continue it (3.4% of comments). Sample comments from the classroom teachers are provided below.

[8] Percentages do not add up to 100% because a single comment may have addressed more than one theme.

What are the strengths of your school's literacy program?

> "It provides differentiated instruction and meets students' individual needs. It allows students to share with peers and engage in small group discussion which promotes interest on their part."
>
> "It covers the standards and there isn't a question of if you covered it or not. I also like the one on one opportunities presented in the daily literacy program."

What areas of your school's literacy program could be improved?

> "We could use more leveled readers. The amount of guided reading we are teaching sometimes makes it challenging to find multiple copies at the A, B and C levels."
>
> "We need more staff to service all the children who need [intervention services]. We need more consistent communication between administration and classroom teachers in the form of shared collaboration. Teachers also need to be able to give their input on whom should be chosen for LLI services. It should not be based ONLY on a one time score."

Why should your school continue or not continue with the current literacy program?

> "Because it works!"
>
> "I feel as though my school should definitely continue with our current literacy program because it provides excellent instruction for students at their academic level. Each range of students is receiving exactly what they need to achieve."

Focus Groups

Two focus groups – one with the LLI teachers and one with the on-site researchers responsible for conducting the LLIOT observations and DIBELS assessments – were conducted in each district at the end of the study in May and June 2010. Approximately 25 LLI teachers and 9 on-site researchers across both districts participated in the focus groups. For both the LLI teachers and the on-site researchers, the focus groups were qualitatively coded into themes and categories from each of the two focus group sessions to summarize the participants' perceptions of LLI. Each of the 166 LLI teacher responses and the 84 on-site researcher responses was represented in a category or individually under its thematic area; no responses were omitted from qualitative analysis. Then, the responses in each category were summed across both focus group sessions for each group (LLI teachers and on-site researchers) to determine the most frequent responses overall for each theme. The following paragraphs summarize the results.

LLI Teachers

An interview protocol with five questions was utilized in the focus groups with LLI teachers. The participants discussed their perceptions related to their general view of LLI, logistical issues they encountered with implementing the system throughout the school year, LLI's strengths and areas for improvement, and suggestions to improve the LLI online data management system, which was piloted by the publisher during the current study. Results are summarized by question below.

General View of LLI

Overall, most of the LLI teachers who participated in the focus groups stated that they liked LLI and felt that it was beneficial to their students. Many of these teachers were particularly impressed with the books used in the system. A few teachers reported liking LLI because it provides them with everything that they need to teach, while others commented that it is fast-paced, engaging to students, and well-organized with an effective format and structure. Some teachers responded that LLI needs more work, describing it as labor intensive and time consuming. Specific areas of concern for these teachers included the pace of the system, which they felt was too fast for their lower-level students, and the large amount of information to cover in each lesson (particularly even-numbered lessons). Several teachers also described the school-level variables that they felt needed improvement in order to implement LLI correctly, including the need for more time, more administrative support, more materials, and more parental support. With regard to the last item, teachers recommended a bilingual informational video that could be sent home to parents at the beginning of LLI as well as a policy of removing students with frequent absences from LLI . Finally, a few teachers commented that LLI can only be implemented effectively if the teachers are extremely organized and familiar with the material.

Logistical Issues with Implementation

When asked about the logistical issues related to implementation, most of the LLI teachers mentioned time and/or scheduling, particularly coordinating with the classroom teachers' schedules in order to meet with all of their LLI groups. Several teachers also commented that they could not finish the lesson during the designated 30-minute timeframe, while a few pointed out that interruptions (including assemblies, half days, state assessments, and the fact that LLI teachers were often pulled to assist with non-intervention activities) frequently prevented them from meeting with their groups. A few teachers also stated that they did not have enough planning time to adequately prepare for their LLI groups.

In addition to the scheduling issues, some LLI teachers commented that they experienced resistance from classroom teachers throughout the LLI implementation; however, these respondents speculated that the classroom teachers' resistance was related to this research study rather than LLI itself (i.e., the classroom teachers did not understand that, due to the need for a treatment and control group, some of their LLI-eligible students would not receive LLI right away). Respondents pointed out that they received more support from the classroom teachers when they worked with them on scheduling. A few respondents even reported that some classroom teachers came to observe their LLI groups in order to learn more about LLI; as a result, these classroom teachers became more supportive of LLI.

Other logistical issues that were reported less frequently included grouping students at different levels from different classrooms, becoming familiar with all of the LLI materials, and the large amount of photocopying involved in preparing lessons, as well as a perceived lack of lesson "flow," insufficient practice activities for slower learners, and the need for more parental support.

Strengths of LLI

The LLI instructional materials, particularly the books and take-home books, were most frequently mentioned by the LLI teachers in the focus groups as strengths of LLI. Respondents felt that the books were high-quality (well-written, fun, colorful, etc.), covered a variety of high-interest topics, and were enjoyable to students. Further, they appreciated the take-home books because many of their students did not have books at home, and the take-home books reinforced learning outside of the LLI group. Teachers also liked the Teacher's Guide and the Lesson Resources CD. In addition to the

materials, several teachers also mentioned the guided lessons and the fact that everything was already in place for them to teach, while some mentioned the writing component or the fact that the system includes both a reading and a writing component. A few teachers also cited the 3-to-1 group size, the Reading Records, the repetition and consistency of lessons, the lesson layout, and the ease of implementation or "teacher-friendliness." Finally, other strengths mentioned included concept building, the integration of both phonics and comprehension, the use of Fountas & Pinnell's Continuum of Literacy Learning (2007), the LLI teacher's ability to meet with the same groups every day, and the fact that the system addresses both instructional and independent reading levels.

Areas for Improvement

When asked about the areas of LLI that could be improved, LLI teachers most frequently mentioned the timeframe for the lessons, stating that they could not complete the lesson in 30 minutes and generally required approximately 45 minutes to cover all of the material. Several teachers also stated that there were inconsistencies between materials (e.g., the lesson did not "match" the written materials), that the system was too fast-paced for at-risk students, and that the amount of information presented in each lesson could be overwhelming to students. Several teachers also recommended improving the lesson sequence or "flow" and the word work, which they stated should be consistent and derive directly from the text. Some teachers also disagreed with the examples used for certain concepts in the curriculum, such as using "ear" for the long "e" sound, presenting "bread" and "read" at the same time (i.e., one is an example of the "ea" rule while one is an exception), and including words like "moon" and "spoon" in the same word ladder as words like "book." Further, some teachers suggested providing more skill review so that mastery can be obtained before progressing to the next skill, providing a more specific vocabulary introduction for each book, and reducing the amount of "runoff" paper that is wasted when printing out CD resources (i.e., some pages print with only a few words per page, which can be problematic for teachers in schools that limit their number of photocopies per month). Finally, other recommendations included providing tips to help students who continue to struggle; increasing the size of the ABC, word, and picture cards; and incorporating a handwriting component.

Suggestions for the Online Data Management System

LLI teachers were also asked about their experiences with the new online data management system and any suggestions that they might have for improving the system. Overall, the most frequent complaint was that the system was slow or tended to "freeze," sometimes for hours at a time. Some teachers reported that this problem worsened if more than one teacher at their school attempted to use the system at once. Many teachers also stated that their data would not save or that it would disappear and reappear, particularly in the Notes section. Several teachers also reported difficulty with the individual and group weekly records; however, these problems may have been due to the fact that some teachers did not know to first enter the information in the group weekly record and then edit it for individual students if needed in the individual weekly record. Suggestions for the weekly records included improving the ability to enter reasons for teacher and student unavailability, adding an option for "group unavailable," and applying school holidays to all of a teacher's groups once she enters the holidays for one group. Some teachers also noted that there is not a way to transfer students in the system from one group or LLI teacher to another, which means that the student still appears in his/her original group and that the data regarding the student's progress in the old group and in the new group are not connected. Finally, some teachers complained about the process of obtaining technical support, stating that the publisher's helpline employees were not able to fix the problem or that they had to talk to several employees before finding someone that could help. Teachers also reported that they were not always able to be on their computer when calling technical support (e.g., if they must use the

telephone in their school's front office) and recommended a link to a frequently asked questions page or a live help chat.

Overall, teachers suggested that the use of the online data management system could be improved with on-site, in-person training on the computer, a more user-friendly manual, and additional resources, such as a step-by-step demonstration video or presentation that could be available for download on the publisher's website, a quick-start guide, or a list of "helpful hints." Most of the teachers indicated that the system had a relatively steep learning curve and took some time to "figure out." However, several of these teachers reported that the system worked well and was easy to use once they learned how to use it. A few teachers reported not experiencing any problems with the system. Finally, some teachers commented that they liked the system, particularly the layout and the ability to access individual and group reports.

On-site Researchers

Because the on-site researchers for the study were all retired teachers who had experience teaching in the two school districts in the study, CREP researchers utilized focus groups to solicit their feedback regarding LLI. The on-site researchers were able to provide an objective "outsider's" perspective based on their random observations of the LLI groups. On-site researchers were asked to discuss their perceptions of LLI's strengths and areas for improvement as well as their opinion of LLI in general (i.e., as experienced teachers, what would they think of having LLI in their schools?). Results are summarized by question below.

Strengths of LLI

Most of the on-site researchers cited group management by the LLI teachers as a strength of the LLI implementation in the current study (e.g., the use of prompts, reinforcement, and rapport with the students and the ability to deal effectively with behavior problems). In terms of LLI in general, on-site researchers frequently cited the fact that LLI builds student confidence, is very well-organized, and can be adapted to different teaching styles and real-world learning situations. Some on-site researchers were also impressed with LLI's fast pace, scripted lessons and routines, and 3-to-1 group size. Other strengths mentioned by the on-site researchers included the books, the level of training received by the LLI teachers, and the range of students that can benefit from LLI (e.g., older students as well as early childhood students and students with special needs or behavior problems).

Areas for Improvement

When asked what areas of LLI could be improved, on-site researchers most frequently responded that the Reading Records were too long and took up too much time during the even-numbered lessons. They recommended establishing a time or word limit for the Reading Record, providing more activities for other students to complete during the assessment time, or even allowing the Reading Record to be done on a separate day if possible. Some on-site researchers also felt that the Reading Record should be used as a general indicator of progress and that the LLI teacher should be able to decide how much of the record to complete for each student.

In addition to their comments about the Reading Records, several on-site researchers echoed the LLI teachers' concerns that the lessons could not be completed in 30 minutes and that the system was too fast-paced for slower learners. A few on-site researchers also felt that the group format could be distracting for younger students, particularly when doing choral reading. Other areas for improvement suggested by the on-site researchers included a perceived overemphasis on comprehension, a belief that the skills presented in the system are too advanced for struggling students,

the difficulty in sharing materials between groups, and the need for more writing and a stronger parental component. On-site researchers also recommended better communication between LLI teachers and classroom teachers in order to streamline the process of getting students to their LLI groups on time as well as to improve the classroom connection component of the curriculum.

Overall Opinion of LLI

When asked what they would think as experienced teachers about having LLI in their schools (i.e., the benefits and drawbacks of implementing such a system), the majority of on-site researchers in the focus groups were positively supportive of LLI and indicated that they would like to have this system if they were still teaching. On-site researchers generally saw LLI as a systematic approach to literacy intervention that covers all of the necessary components, is high-interest for students, and can be effective with a range of early childhood students. On-site researchers commented that they liked the system's leveled texts, small group format, instructional techniques (including the emphasis on phonics and the reading/writing connection) and the use of repetition, as well as the LLI teacher's ability to see growth in her students' literacy over time. Some on-site researchers also commented that LLI is similar to the Reading Recovery program, which would make it easier for districts already implementing Reading Recovery to adopt LLI.

Some on-site researchers also commented on the potential drawbacks of implementing LLI. These on-site researchers pointed out that the system can be expensive; requires extensive training; involves a good deal of documentation; and would not be implemented as effectively by less experienced teachers. Finally, some on-site researchers provided a "conditional" answer to this focus group question, stating that their opinion of LLI would depend on who was implementing it. Specifically, these on-site researchers perceived that LLI would only be effective if it was implemented in small groups by an experienced early childhood teacher with extensive training in LLI. These on-site researchers further emphasized that "every piece of the program must be in place" for the most successful implementation of LLI.

Conclusions

1. What progress in literacy do students who receive LLI make compared to students who receive only regular classroom literacy instruction?

Across the three grade levels, the current study found that LLI positively impacts K-2 student literacy achievement in rural and suburban settings. Further, we determined that LLI is effective with ELL students, students with a special education designation, and minority students in both rural and suburban settings. Finally, the current study showed that LLI is effective with economically disadvantaged children in both rural and suburban settings.

This study found robust effects on the LLI Benchmarks across all grade levels for students who received LLI. Across the three grade levels, students in LLI achieved between 1 ½ benchmark levels up to almost 5 ½ benchmark levels, while students who did not receive LLI achieved between less than 1 benchmark level to 3 benchmark levels.

Further, these effects were particularly strong for various subgroups (e.g, ethnicity, special education or ELL status) within each grade level. For kindergarten, significant effects were found, compared to the control group, for African American students, Hispanic students, and ELL students on the LLI Benchmarks, with all three subgroups finishing closer to grade level (i.e., Level C) than their counterparts who finished at or below Level A. 1st grade African American and Hispanic students in the treatment group also showed more gains than their counterparts in the control group. In 2nd grade, strong, educationally meaningful effects were found for African American and Hispanic LLI students. 2nd grade African American LLI students finished at the highest level overall, closely followed by the Hispanic LLI students.

Additionally, effects found with the DIBELS measures of reading fluency provided corroboration of the results with the LLI Benchmarks. In kindergarten, students in LLI showed significant gains on subtests of the DIBELS as compared to those who did not have LLI. In particular, for phoneme segmentation fluency, ELL students in the treatment group outperformed ELL students in the control group, *as well as* non-ELL students in both the treatment and control groups. In 1st grade, LLI students significantly exceeded the control group on 3 of 4 subtests: nonsense word fluency, letter naming fluency, and oral reading fluency. Finally, on the nonsense word fluency subtest, 1st grade Hispanic students in the treatment group outperformed their counterparts in the control group.

Taken together, all of the student achievement results provide strong evidence that students who are eligible for and participate in LLI make significant progress in literacy compared to students who are eligible to receive LLI and only receive regular classroom literacy instruction.

2. Was LLI implemented with fidelity to the developers' model?

Overall, the observation results from the current study suggest that LLI was implemented with a high degree of fidelity to design across both districts. The majority of lesson components received high fidelity ratings in most of the observations that were conducted. Further, qualitative feedback from the LLI teachers suggests that two of the lesson components receiving lower fidelity ratings – the classroom and home connections – may have been implemented less frequently due to lack of cooperation or support from parents and lack of familiarity with LLI by the classroom teachers.

Additionally, observation results revealed that LLI implementation was consistent across the year, with high fidelity scores received at both time points when the observations were conducted. The only change in implementation over the year occurred in first and second grade, where "Learning Environment" scores significantly improved from the first observation to the second. Because this scale of the observation tool measured such factors as lesson organization and pacing, the availability of instructional materials, and an overall rating of whether the lesson was delivered as designed, this finding may have been due to a "practice effect" in which teachers became more familiar and comfortable with the materials and procedures over time.

Finally, the LLI attendance records from the current study revealed that, on average, students received less than the model's recommended number of instructional days (i.e., approximately 70 days instead of 90 for 1st and 2nd grade, and approximately 40 days instead of 70 for kindergarten). However, despite receiving less than the recommended amount of instruction, students in all three grade levels made significant progress in their literacy achievement. This finding suggests that LLI can still be effective during a relatively shorter timeframe, which may be valuable to districts with a large number of students to serve or limited time in which to implement early literacy interventions.

3. What were LLI teachers' perceptions of LLI and its impact on their students' literacy?

Overall, the LLI teachers in the current study supported LLI and believed it had a positive impact on their students' literacy. LLI teachers indicated that they had a good understanding of LLI; received support in implementing LLI from their district, school administration, and other school staff; and perceived a positive impact on student achievement and student attitudes towards literacy. LLI teachers also reported a positive impact of LLI on their reading instruction, particularly their understanding of the role of comprehension and phonics/phonemic awareness in the reading process and the relationship of leveled texts to successful reading. LLI teachers were extremely positive about the materials and resources, particularly the leveled texts, which they described as high-quality, high-interest, and engaging to students. LLI teachers also liked the small group format of the lessons as well as the well-organized, guided lesson structure. However, a number of LLI teachers in the current study thought the LLI lessons contained too much information to adequately complete a lesson during the designated 30-minute timeframe. Many LLI teachers also thought the system was too fast-paced for their lower-level students and that there were inconsistencies in materials which made implementation of certain lessons more difficult.

In addition to the LLI teachers, a small number of classroom teachers with students in the current study provided feedback on their perceptions of the LLI system. Most of these teachers were positive about LLI and noticed that their students' literacy in the classroom improved after receiving LLI, with one classroom teacher even commenting, "…I believe that children that struggle would give up hope in the realm of reading without the LLI program."

Recommendations

Altogether, the results from this evaluation allow us to conclude that the LLI system positively impacts students' literacy skills. These results also suggest that continued implementation of LLI would be beneficial in both Tift County Schools and the Enlarged City School District of Middletown. While the long-term impacts of LLI have yet to be determined, the positive results found in this evaluation suggest that additional benefits may be seen with the continuation of LLI. This evaluation provided a randomized controlled trial and efficacy study for the LLI system as well as offered an opportunity for research-based recommendations that may enhance the system, future research, and ultimately student achievement. From this evaluation, CREP proposes the following recommendations with regard to LLI and its implementation in schools:

- When possible, schools should begin kindergarten instruction in LLI as soon as possible in order to provide the recommended amount of instruction (i.e., 14 weeks) for kindergarten students.
- Professional development for building principals and central office supervisory staff, although not measured in this study, surfaced as being critical to the implementation.
- Likewise, regular classroom teacher involvement and professional development to familiarize them with LLI and its features also appears to influence the quality of implementation.
- LLI teacher professional development should be ongoing with at least a refresher training to supplement and resolve any district-specific issues.
- Providing scenarios or examples of how prior adopters have developed schedules that allow for full implementation of the 30-minutes-a-day, five-days-a-week instructional pattern would be helpful to school districts who are new adopters of LLI.
- Suggestions and recommendations of how LLI teachers might plan and organize their LLI sessions so they can accomplish the instructional goals in a typical 30-minute session would benefit prior and new adopters of LLI.
- Additional suggestions from the authors about how best to instruct LLI groups whose members are not at the same level or who have members progressing at a slower rate would be helpful.
- Providing some type of video for parents of the LLI students could not only explain the system but could provide clips of how they should be working with their child. This is particularly important for the parents of ELL children and the parents of economically disadvantaged children.
- A careful review of all materials and resources is recommended to ensure consistency and accuracy throughout the system.
- There is a great need to conduct a similar study in at least one major urban district.
- Future research of LLI should include longitudinal tracking of student reading achievement to look at the long-term impact of LLI beyond one school year.
- The LLI benchmarking system would benefit from additional systematic comparisons with other nationally recognized literacy assessments.

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.

Clay, M. (1991). *Becoming literate: The construction of inner control*. Portsmouth, NH: Heinemann.

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). *Research base for Leveled Literacy Intervention*. Portsmouth, NH: Heinemann. Retrieved August 29, 2008, from http://www.heinemann.com/fountasandpinnell/research/LLIResearchBase.pdf

Fountas, I. C., & Pinnell, G. S. (2009). *Leveled Literacy Intervention*. 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.

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.

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.

National Research Council. (1998). *Preventing reading difficulties in young children.* Washington, DC: National Academy Press.

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.

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.

Wanzek, J., & Vaughn, S. (2007). Research-based implications from extensive early reading interventions. *School Psychology Review, 36*(4), 541-561.