Fountas & Pinnell

HEINEMANN

Leveled Literacy Intervention
Research and Data Collection Project
2010–2011

Lindsay B. Demers, Ph.D., Data Collection Manager
(working under the guidance of Irene C. Fountas and Gay Su Pinnell)

Heinemann

DEDICATED TO TEACHERS

HEINEMANN
Leveled Literacy Intervention (LLI)
Research and Data Collection Project
2010-2011

Lindsay B. Demers, Ph.D., Data Collection Manager
(working under the guidance of Irene C. Fountas and Gay Su Pinnell)

Full Report

This report describes the results of a correlational study of *Leveled Literacy Intervention (LLI)* that took place during the 2010–11 school year. The analyses reported here, commissioned by Heinemann, a publisher of professional resources and a provider of educational services for teachers, were conducted by Lindsay B. Demers, Ph.D.

What Is *Leveled Literacy Intervention*?

LLI is short-term, small-group supplementary intervention curriculum for struggling readers in grades K–8+. It was developed by Irene Fountas and Gay Su Pinnell and is built around their F&P Text Level Gradient™, which designates leveled texts from A (the easiest) through Z (the most difficult). The objective of *LLI* is to help students reduce the gap between their current instructional reading level and their expected instructional reading level.

LLI is described as a *system* because it includes reading, writing, and word study, each used symbiotically and methodically in connection with increasingly difficult texts. Three *LLI* systems have been published at the time of this study:

- Orange, recommended for kindergarten (levels A to C);
- Green, recommended for grade 1 (levels A to J); and
- Blue, recommended for grade 2 (levels C to N).

The K–2 *LLI* systems provide 14 to 20 weeks of daily, intensive extra reading lessons (beyond general classroom instruction); decisions about when students enter and leave the program are flexible. Students enrolled in *LLI* participate in daily (five times a week) 30-minute literacy lessons taught by an *LLI* trained teacher. Recommended group size is three students with one teacher, which allows for close observation and differentiated response to student strengths and needs. Because of restrictions or limited resources, some school districts may alter these requirements, but the goal of one teacher for three students and five lessons a week is highly recommended for the best outcomes.

The intended term of *LLI* K–2 ranges from about 14 to 18 weeks, after which the students are released if approximate grade-level expectations are met. At the end of this period, students who have made progress but do not yet meet expected gradient and fluency levels are reevaluated. The evaluation may suggest additional *LLI* lessons or an alternative intervention (such as individual tutoring). Students who are enrolled in the intervention program at the end of the school year, but have not completed it, may continue at the beginning of the next school year. Some students who have completed

the intervention may have achievement levels more than a year below grade level. If they have made steady progress and resources allow, they may continue to participate.

LLI is now being extended for students in grades 3 through 8, reading at levels L through Z. The *LLI Red System* (grade 3, levels L–Z) is available now (delivery in October 2012). The *LLI Gold, Purple* and *Teal* systems (grades 4–8+, levels O–Z) are currently in development. Each system provides 24 lessons at each level, L–Z, with systems overlapping in level for flexibility of use. At the end of the 24 lessons, students will read a full-length novel; four days of additional lessons are provided. Students then participate in four days of optional test preparation.

Student Characteristics

A total of 2,679 students from 114 schools in 11 states participated in this study. The sites included public school districts and individual public schools, charter schools, and private schools. *Table 1* is a breakdown of the geographic areas represented.

Geographical Distribution of *LLI* Students	
Region	**Percentage of Students**
Northeast	26% ($n = 707$)
Midwest	22% ($n = 598$)
South	23% ($n = 629$)
West	28% ($n = 745$)

Table 1.

Of the students, 54% were male and 43% were female (3% were not recorded as either male or female). With regard to ethnicity, 31% of students were reported as White, 37% were reported as Hispanic, 19% were reported as Black, 2% were reported as Asian/Pacific Island, 1% were reported as Native American/Alaskan, 2% were reported as Multiracial, and 1% were reported as Other. (Race/ethnicity data were missing for approximately 6% of participating students.) Thirty-two percent of students ($n = 843$) were English language learners (ELLs); 59% of students ($n = 1588$) received free/reduced-price lunches; 16% of students ($n = 424$) had an individualized education plan (IEP). *Table 2* summarizes the *LLI* student sample by grade, and *Table 3* summarizes the students' race and gender demographics.

Grade Distribution of *LLI* Students					
Kindergarten	**1st Grade**	**2nd Grade**	**3rd Grade**	**4th Grade**	**5th Grade**
12% ($n = 308$)	36% (n = 961)	33% (n = 884)	12% ($n = 328$)	3% (n = 88)	1% (n = 30)

Table 2.

Breakdown of Student Sample by Gender and Race			
	Male	**Female**	**Total**
White	17.8% ($n = 471$)	13.9% ($n = 368$)	31.8% ($n = 839$)
Hispanic	20.6% ($n = 545$)	16.9% ($n = 447$)	37.6 % ($n = 992$)
Black	10.5% ($n = 279$)	9.0% ($n = 239$)	19.6% ($n = 518$)
Asian/Pacific Island	1.4% ($n = 38$)	0.8% ($n = 22$)	2.3% ($n = 60$)
Native Amer./Alaskan	0.5% ($n = 13$)	0.5% ($n = 12$)	0.9% ($n = 25$)
Multiracial	1.5% ($n = 39$)	0.9% ($n = 24$)	2.4% ($n = 63$)
Other	0.5% ($n = 13$)	0.6% ($n = 16$)	1.1% ($n = 29$)
Total	54.3% ($n = 1398$)	43.7% ($n = 1128$)	94% ($n = 2526$)

Note. Gender and race information were missing for 6% of students.

Reading Recovery® (RR) services were provided to 11% of the students ($n = 292$) prior to their participation in *LLI*. Students for whom RR attendance rates were available ($n = 291$) received an average of 73.27 ($SD^1 = 21.18$) RR lessons over 31.92 weeks ($SD = 30.82$). Of those students who participated in RR prior to receiving any *LLI* lessons, 42% ($n = 123$) reached their grade-level equivalent.

Teacher Characteristics

A total of 326 teachers took part in the study. The *LLI* system was new to most teachers (average years teaching *LLI* = 1.65, SD = 1.34, range 1 to 10 years). On average, teachers received 24.63 hours of training (SD = 11.86, range 0 to 60 hours). See *Table 4*.

Teacher *LLI* Training Sources	
Training Source	**Number of Teachers**
Heinemann	24% ($n = 82$)
Lesley University	4% ($n = 12$)
Ohio State University	0
District	10% ($n = 34$)
***LLI* Materials (Self-Taught)**	2% ($n = 8$)
District and *LLI* Materials	<1% ($n = 1$)
Other	50% ($n = 172$)

[1] The standard deviation *(SD)* of a mean *(M)* is a measure of variability around the mean. The larger the standard deviation, the greater the spread of values in the data around the mean.

Implementation

With *LLI* lessons, it is intended that one teacher will work with a small group of three students five days per week. However, in this study, delivery varied because of school and district limitations. Overall, the *LLI* groups contained between one and six students ($M = 3.13$, $SD = 0.65$). The number of lessons per week for students ranged from 1 to 9 ($M = 4.44$, $SD = 0.79$), and the number of weeks in the *LLI* program ranged from 1 to 46 ($M = 18.64$ weeks, $SD = 8.49$). Sixty-three percent of the sample ($n = 1,351$) received the full sequence of *LLI* lessons, according to school administrators. The most common reason students did not complete the entire *LLI* curriculum was because the school year ended while it was being implemented (18%, $n = 377$). See *Table 5*.

Mean (Standard Deviation) *LLI* Implementation Characteristics by Grade				
	Average Group Size	Average No. of Weeks of *LLI*	Average No. of *LLI* Lessons Per Week	Average No. of Total *LLI* Lessons
Kindergarten	3.12 (0.57)	12.95 (4.69)	4.26 (0.95)	41.72 (19.87)
Grade 1	3.08 (0.65)	19.51 (8.89)	4.45 (0.87)	65.13 (32.95)
Grade 2	3.09 (0.63)	19.49 (8.21)	4.47 (0.71)	64.42 (30.52)
Grade 3	3.33 (0.77)	18.28 (8.42)	4.40 (0.64)	55.15 (27.45)
Grade 4	3.44 (0.63)	22.02 (7.91)	4.55 (0.50)	67.76 (27.59)
Grade 5	3.33 (1.06)	16.57 (8.57)	4.71 (0.46)	48.57 (40.10)
Overall	3.12 (0.66)	18.53 (8.47)	4.43 (0.79)	60.57 (31.04)

Table 5.

Assessing Student Reading Gains

Because this study does not include data from a control group, no claims about causation regarding participation in *LLI* can be made, nor can the impact of *LLI* be teased apart from the impact of the regular schooling students received during the period of the study. However, the collected data allows us to compare students' observed reading gains with their expected reading gains. The relationship between *LLI* dosage (number of lessons) and performance on the exit assessment can also be examined.

Observed Reading Gain

Observed student reading gain was determined by looking at students' pre- and post-*LLI* instructional reading levels (ranging from pre-A to Z) based on teacher-reported results from the *Fountas & Pinnell Benchmark Assessment System* or from student reading records. In some cases, levels were determined by converting the results from a different assessment (e.g., the DRA) to Fountas & Pinnell levels. Those levels were then converted to time equivalent (TE) scores (see *Table 6*). TE scores represent the number of months students, with typical progression, should have completed at a given point during their schooling according to the Fountas & Pinnell levels and assuming a ten-month school year. If a level is expected to be reached over a period of more than one month, the TE score is the middle point of the expected band of time. For example, students are expected to be at level C during months eight, nine, and ten during kindergarten; the TE score for level C is 9. To determine students' gain, a difference score was calculated by subtracting a student's TE score when he or she entered *LLI* from his or her TE score when the student exited *LLI*. For example, a student who enters *LLI* at level A and exits at level C has progressed from the expected level at the fourth month of kindergarten to the expected level at the ninth month of kindergarten; this student's reading gain is five months.

Fountas & Pinnell Instructional Level Expectations for Reading and Time Equivalent Scores										
	Month 1	Month 2	Month 3	Month 4	Month 5	Month 6	Month 7	Month 8	Month 9	Month 10
Kindergarten	Pre-A TE=2	Pre-A TE=2	Pre-A TE=2	A TE=4	A/B TE=5	B TE=6.5	B TE=6.5	C TE=9	C TE=9	C TE=9
1st Grade	C/D TE=11	D TE=12	E TE=13	E/F TE=14	F TE=15	G TE=16	G/H TE=17	H TE=18	I TE=19.5	I TE=19.5
2nd Grade	I/J TE=21	J TE=22.5	J TE=22.5	J/K TE=24	K TE=25	K/L TE=26	L TE=27.5	L TE=27.5	M TE=29.5	M TE=29.5
3rd Grade	M/N TE=31	N TE=33	N TE=33	N TE=33	O TE=36	O TE=36	O TE=36	P TE=39	P TE=39	P TE=39
4th Grade	P/Q TE=41	Q TE=43	Q TE=43	Q TE=43	R TE=46	R TE=46	R TE=46	S TE=49	S TE=49	S TE=49
5th Grade	S/T TE=51	T TE=53	T TE=53	T TE=53	U TE=56	U TE=56	U TE=56	V TE=59	V TE=59	V TE=59

Table 6.

Of the 2,460 students from whom pre- and post-data were collected, 2,350 (about 95%) exhibited growth from pre- to post-assessment. Eleven students' scores decreased from pre- to post-assessment, and 99 students' scores did not change from pre- to post-assessment. The average gain score was just under 9 months (M = 8.72 months, SD = 5.38; median = 8 months) but varied a great deal from –5 months (in the case of students whose scores decreased) to 40 months. The average reading gain of about 9 months is particularly noteworthy because the average time LLI was administered was approximately 4.5 months. Therefore, students who participated in LLI made twice the progress that would be expected within that time frame (i.e., an average reading gain of nearly 9 months in just 4.5 months' time). Average TE entry scores, exit scores, and change scores, by grade, are provided in *Table 7*.

Mean (Standard Deviation) of Student Assessment Scores and Growth Rates			
Grade	TE Entry Score	TE Exit Score	Gain in Months
Kindergarten (n = 288)	3.41 (1.48)	7.47 (2.39)	4.07 (2.38)
1st Grade (n = 878)	8.21 (4.29)	16.89 (4.26)	8.67 (5.04)
2nd Grade (n = 826)	15.73 (5.00)	25.58 (5.53)	9.85 (5.55)
3rd Grade (n = 318)	20.32 (6.72)	29.75 (5.73)	9.43 (5.42)
4th Grade (n = 85)	16.19 (4.68)	27.32 (6.03)	11.12 (5.41)
5th Grade (n = 30)	20.70 (5.90)	30.70 (4.51)	10.00 (5.45)
Overall (n = 2,425)	12.22 (7.20)	20.95 (8.43)	8.72 (5.37)

Note. Refer to Table 6 to convert these numerical values into F & P reading levels.

Table 7.

Expected Reading Gain

Expected student reading gain was determined by calculating what the TE scores for the entrance and exit assessments should be, given the point in the school year students were when they entered and exited *LLI*. For example, assuming that school began in September (month 1), if a student entered *LLI* during October of kindergarten, her expected TE entry score would be 2. If she exited *LLI* during February of kindergarten, her expected TE exit score would be 6.5. Expected reading gain scores were calculated by subtracting the expected entry score from the expected exit score. Thus, the student who started *LLI* in October of kindergarten and who exited *LLI* in February of kindergarten would have an expected gain of 4.5 months.

Observed Versus Expected Reading Gain

Comparing the observed and expected assessment scores at the beginning and end of *LLI* gives an estimate of the extent students' reading skills were behind what would be expected of a student at their educational level before and after participation in *LLI*. This comparison also reveals whether the discrepancy between observed and expected assessment scores decreased after the students participated in *LLI*. (For technical details about the statistical analyses employed here, see the appendix.)

- On average, students' observed reading scores were seven months behind what would be expected (average discrepancy between expected and observed reading scores at entry = 7.29 months, $SD = 6.68$). The magnitude of the discrepancy between observed and expected scores on the entry assessments was statistically significant, meaning the differences observed were not likely due to chance (see *Figure 1*).

Discrepancy Between Observed and Expected Student Entry Assessment Scores ($n = 2,524$)

5%   3%

92%

- Scored Above Expected Level on Entry Assessment
- Scored at Expected Level on Entry Assessment
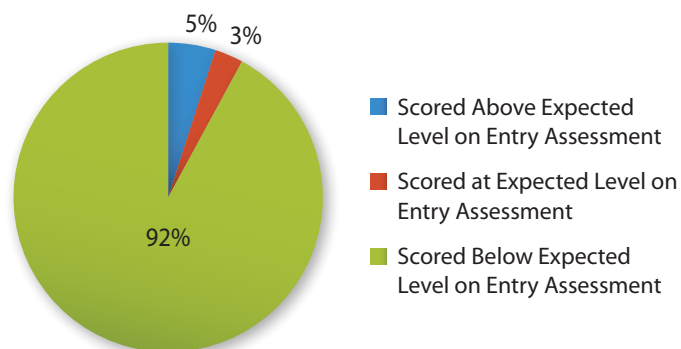- Scored Below Expected Level on Entry Assessment

Figure 1.

- Upon exiting *LLI*, students' reading scores were just three months behind what would be expected (average discrepancy between expected and observed reading scores = 3.34 months, $SD = 6.87$). The magnitude of the discrepancy between observed and expected scores on the entry assessments was statistically significant, meaning the differences observed were not likely due to chance (see *Figure 2*).

Discrepancy Between Observed and Expected Student Exit Assessment Scores (*n* = 2,401)

- Scored Above Expected Level on Exit Assessment
- Scored at Expected Level on Exit Assessment
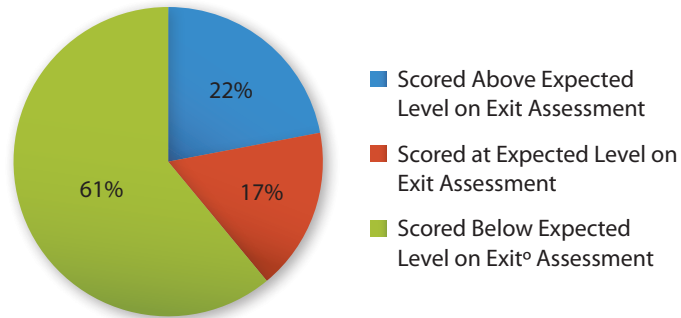- Scored Below Expected Level on Exitº Assessment

22%

61%

17%

Figure 2.

- Although many students scored below the expected level on both their entry and exit assessments, the magnitude of the discrepancy between expected and observed scores at entry was significantly greater than the magnitude of the discrepancy between expected and observed scores at exit. Therefore, although on average students never reached their expected achievement level, the magnitude of the difference between observed and expected reading scores was smaller after participation in *LLI* (see *Figure 3*).

Difference: Discrepancy Between Observed and Expected Student Entry Assessment Scores and Discrepancy Between Observed and Expected Student Exit Scores (*n* = 2,353)
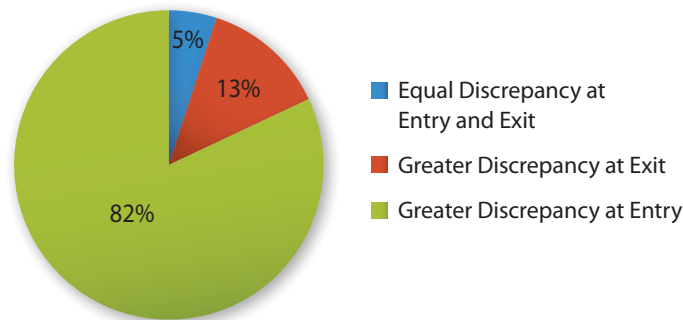

- Equal Discrepancy at Entry and Exit
- Greater Discrepancy at Exit
- Greater Discrepancy at Entry

5%

13%

82%

Figure 3.

Relationship Between *LLI* Dosage and Exit Score

As noted above, not all students received *LLI* as it was intended to be delivered. Across grade levels, the average number of lessons was about 60, though this ranged widely from 2 to 185. One advantage of having this wide variability in *LLI* dosage is that it creates an opportunity to look at the relationship between total number of *LLI* lessons and performance on the exit assessment. (Technical details about the statistical analyses used here are provided in the appendix.)

- Even after taking into account students' entry assessment scores, for each additional *LLI* lesson students received, their exit assessment scores increased by one-tenth of a point (see *Figure 4*). This number was statistically significant, which means that the gain in exit scores observed for each additional *LLI* lesson was not likely due to chance.

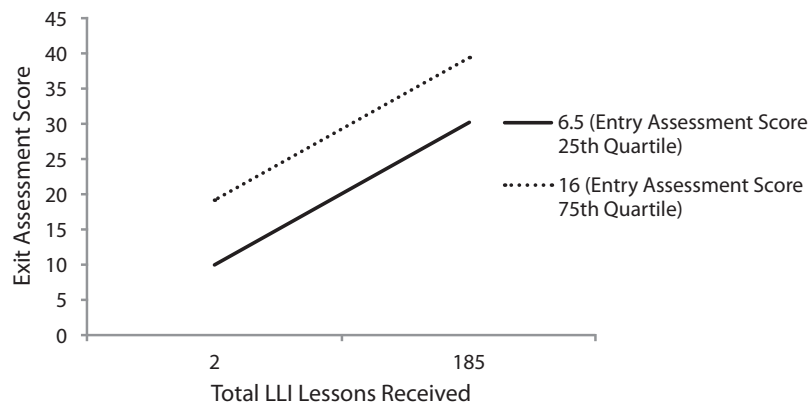## The Relationship Between Total *LLI* Lessons and Exit Assessment Score

Figure 4.

Limitations

The findings reported here suggest that participation in *LLI* is strongly associated with higher scores on student exit assessments. However, it is important to bear in mind that no data were collected from a control group (i.e., a group that received no intervention) so no claims can be made about the causal effects of participating in *LLI*, nor can it be determined whether the students who participated in *LLI* scored above what would be expected simply as a result of being in school for several months. The number of months in school and the number of *LLI* lessons were highly correlated ($r = 0.85$). Thus, the unique impact of *LLI* in these data cannot be teased apart from the impact of the typical schooling each child would have received. In order to address these concerns, experimental data with a control group must be collected.

Appendix: Statistical Analyses

To examine the relationship between observed and expected TE scores, nonparametric sign tests were used. These tests were used in place of paired-sample t tests because the distributions of these variables were nonnormal. The first sign test compared observed entry score with expected entry score. Of the total 2,524 students who were included in this analysis, only 113 had observed entry levels that surpassed their expected reading levels; 77 had observed entry levels that were in line with their expected entry level scores; and 2,334 had observed entry scores that were below their expected entry level. This test was highly significant ($p<0.001$), showing that the number of students who were below their expected reading level was much greater than would be expected by chance.

The second sign test compared observed exit score with expected exit score. Of the 2,401 students included in this analysis, 525 scored above their expected exit level; 411 scored at their expected exit level; and 1,465 scored below the expected level. This test was also highly significant ($p<0.001$), showing that the number of students who were below their expected reading level on the exit assessment was greater than would be expected by chance.

The third sign test compared the discrepancy between observed and expected assessment scores at entry and exit to see whether the magnitude of the discrepancy decreased after participation in *LLI*. Of the 2,353 students included in this analysis, 1,928 showed a greater discrepancy between their observed and expected assessment scores at entry compared with the discrepancy at exit; 119 showed equal discrepancy at both time points; and 306 showed a greater discrepancy between observed and expected assessment scores at exit compared with entry. This test was highly significant ($p<0.001$),

showing that the number of students whose discrepancy scores decreased after participation in *LLI* was greater than would be expected by chance.

Hierarchical linear modeling (Raudenbush & Bryk 2002) was used to examine the relationship between total *LLI* lessons and exit assessment score, in order to account for the variance shared by students who had the same *LLI* instructor. Variance estimates from a completely unconditional model showed that over half (53%) of the variance in students' exit scores was due to teacher characteristics, while 47% of the variance in students' exit scores was due to student characteristics. Initially, a three-level model was fit to the data that included school as the highest-level unit of analysis. However, preliminary analyses showed that there was not statistically significant nesting at the school level (only about 3% of the variance was due to school characteristics); thus the simpler, two-level model was retained for subsequent analyses.

To start, theoretically relevant student-level control variables (entry assessment score, IEP status, free or reduced-price lunch status, and ELL status) were included in the level-1 model one at a time. Interaction terms between entry assessment score and *LLI* dosage were also tested, given that a student's tenure and/or success in *LLI* may vary as a function of his/her baseline achievement. Covariates were retained in the model if the coefficient associated with each predictor reached significance, and also if inclusion of the predictor significantly improved the fit of the model. Subsequent to finding the best-fitting control model to the data, total *LLI* lessons was included as a test predictor.

The best-fitting model that was retained for interpretation included student entry assessment scores and the number of *LLI* lessons received as level-1 predictors. These two variables did not interact. Although it may initially seem surprising that entry assessment score was the only student-level predictor, it is important to bear in mind that any differences in student exit scores that may have been attributable to ELL status or IEP status, for example, were likely also reflected in the student entry scores, which may explain why these variables did not explain a significant amount of variance above and beyond entry assessment scores (Ballou, Sanders & Wright 2004). In the level-1 model, entry assessment scores explained approximately 35% of the variance observed in exit scores. Total number of *LLI* lessons explained approximately 12% of the variance observed in exit scores.

Fixed Effect	Coefficient	Std. Error	t	p
For Intercept, β_0 γ_{00}	3.69	0.61	6.05	<0.001
For Entry Score Slope, β_1 γ_{10}	0.95	0.03	31.25	<0.001
For Total *LLI* Lessons Slope, β_2 γ_{20}	0.10	0.01	14.62	<0.001
Random Effect	**Variance Component**	**Std. Deviation**	x^2	p
Intercept, u_0	14.76	3.84	154.91	<0.001
Entry Score, u_1	0.04	0.21	141.91	<0.001
Total *LLI* Lessons, u_2	0.002	0.05	169.01	<0.001
s_2	8.48	2.91		

References

Ballou, D., W. Sanders & P. Wright. 2004. "Controlling for Student Background in Value-Added Assessment of Teachers." *Journal of Educational and Behavioral Statistics* 29(1): 37–65.

Raudenbush, S.W. & A.S. Bryk. 2002. *Hierarchical Linear Models: Applications and Data Analysis Methods.* Thousand Oaks, CA: Sage.

Fountas & Pinnell

Leveled Literacy Intervention

RESEARCH AND DATA COLLECTION PROJECT
2010–2011

Heinemann

DEDICATED TO TEACHERS

**To order or more information
www.heinemann.com
Phone: 800.225.5800
Fax: 877.231.6980**