

Language, Literacy & Vocabulary!



Evaluation Study of Language, Literacy, & Vocabulary!

Spring 2007

submitted to:

Barbara Quincer Coulter

Director of Research and Professional Development

National Geographic School Publishing

Evanston, IL 60201

847-425-6320

bquincer@NGSP.com



Submitted by:

Dr. Mary McNabb, Director

Learning Gauge, Inc.

P.O. Box 1238

St. Charles, IL 60174

630-606-3057

mlmcnabb@learningauge.org

Table of Contents

Executive Summary	3
Purpose of the LLV Study	4
Methodology	4
Statistical Results	5
LLV Classroom Adoption Practices	12
Conclusions and Recommendations	13
Appendices	
Appendix A: Validation of the criterion-referenced test	14
Appendix B: Descriptive statistics by learner type	29
Appendix C: Descriptive statistics by grade level	32
Appendix D: Group frequency charts by grade level	34
Appendix E: Histograms	40
Appendix F: Exist Interview Questions	43
Appendix G: WOL Survey	44
Appendix H: RE Survey	47
Appendix I: Classroom Observation Protocol	50

Executive Summary

Purpose

A validation study of the *Language, Literacy, & Vocabulary!* program was conducted by Learning Gauge, Inc. in partnership with National Geographic School Publishing, during the 2006-07 school year. The program evaluation examined classroom adoption approaches used by participating teachers and the subsequent impact of the *Language, Literacy, & Vocabulary!* (LLV) curriculum on students' content area literacy development. The LLV program provides curriculum resources for vocabulary and comprehension development through study of academic content. The LLV materials are designed for multi-level instruction of at-risk readers, English language learners, and other special needs learners. The LLV materials can also be used with grade level readers in a variety of instructional situations.

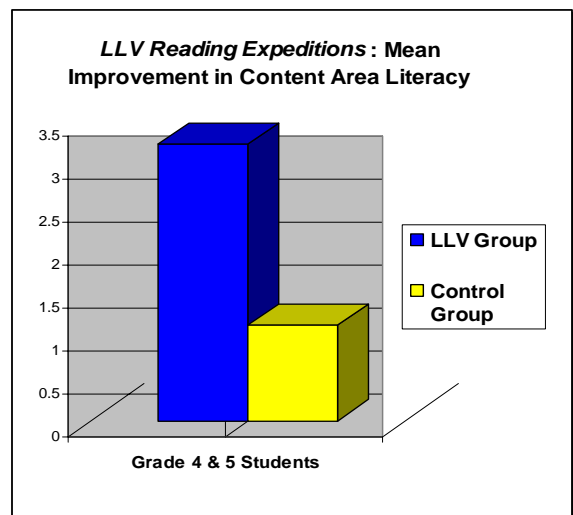
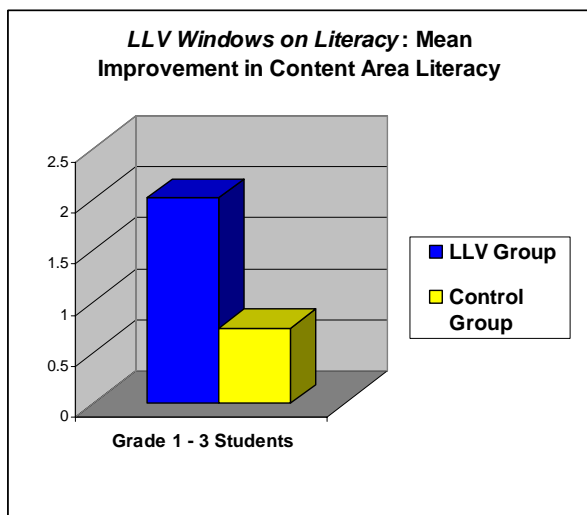
Methodology

The study involved 442 students from three public elementary schools. One school was located in the Southwest and two in the Midwest. For the paired samples t-test analysis, there were 253 students in treatment groups and 189 students in the control groups for grades 1 through 5. Of the total 253 treatment students who completed the study, 12% were English language learners, 24% were at-risk readers in English only, and 64% were grade-level readers. Student groups came from intact classrooms. Pre- and post-test data from students provides insight into how effectively the *Language, Literacy & Vocabulary!* curriculum improves students' content area literacy. Qualitative data from pre-and post-surveys, classroom observations and interviews with treatment and control teachers supplemented the student criterion-referenced test data.

Results

Findings indicate that student groups whose teachers implemented the full range of LLV curriculum and assessment materials with students on a daily basis for at least 5 weeks had statistically significant learning gains across the three learner types: grade-level readers, at-risk readers, and English language learners.

The data suggests that teachers' focus on the LLV comprehension strategies versus content learning has a positive impact on students' reading comprehension of academic texts. The type of learner also impacts reading achievement from opportunities to learn with the LLV curriculum. At-risk readers and English language learners in treatment groups out performed their control group counterparts in every grade level except for the at-risk students in 1st grade. Further study of students with special needs is warranted due to their low participation rates in this study. Grade-level readers in all of the treatment classrooms performed significantly better than grade-level readers in the control groups.



Purpose of the LLV Study

A validation study of the *Language, Literacy, & Vocabulary!* program was conducted by Learning Gauge, Inc. in partnership with National Geographic School Publishing, during the 2006-07 school year. The program evaluation examined classroom adoption approaches used by participating teachers and the subsequent impact of the *Language, Literacy, & Vocabulary!* (LLV) curriculum on students' content area literacy development. The LLV program provides curriculum resources for vocabulary and comprehension development through study of academic content. The LLV materials are designed for multi-level instruction of at-risk readers, English language learners, and other special needs learners. The LLV materials can also be used with grade level readers in a variety of instructional situations.

Methodology

The evaluation for the *Language, Literacy & Vocabulary!* curriculum involved a two-fold design. First, a validation study of the criterion-referenced tests developed to measure students' academic vocabulary and reading comprehension strategy use was conducted. For details about the validity and reliability of the measurement used in this study refer to Appendix A. Second, the pre-and post-test results from the validated measurement were statically analyzed. Additional data were gathered from treatment group teachers about their classroom adoption practices with regard to the *Language, Literacy & Vocabulary!* curriculum. These data allowed for analysis of teachers' program implementation logs, a pre- and post-implementation survey, and exit interviews with each teacher from the control and treatment classrooms. A contact person at each school completed one or two classroom observations for each treatment and control classroom to document the pedagogical practices used in participating classrooms. Control group teachers also were interviewed at the end of the study to gather information about their classroom practices. Pre-and post-test data from students in the treatment groups provides insight into how their teachers' implementation of *Language, Literacy & Vocabulary!* impacted students' content area literacy.

The study involved 442 students from three public elementary schools. Originally, there were four schools recruited for the study, two in the Southwest and two in the Midwest, which were paired for essential demographic characteristics. However, one of the southwest schools dropped out of the study in early fall 2006. There was a high mobility rate in the remaining Southwest school, which had a 95% Hispanic student population. Thus, the low number of English language learners in this study limits statistical claims about that specific subgroup of students. For the paired samples t-test analysis, there were 253 students in treatment groups and 189 students in the control groups for grades 1 through 5. Of the total 253 treatment students who completed the study, 29 were English language learners (12%), 60 were at-risk readers in English only (24%), and 164 were grade level readers (64%). Student groups came from intact classrooms.

At the beginning and conclusion of the study, students were administered a grade level criterion-referenced test. Students also completed formal and informal assessments that were part of the *Language, Literacy & Vocabulary!* curriculum, although assessment usage varied among teacher groups. Qualitative data from pre-and post-surveys, classroom observations and interviews with teachers was collected to supplement the student test results. These data indicated that teachers in the treatment classrooms implemented the LLV curriculum in a variety of ways. The data suggest that teachers' instructional use of the curriculum has a strong impact on student learning. Teachers in the treatment classrooms were assigned to teach five weekly units and received a half-day of professional development orientation to the curriculum. Actual implementation among treatment

teachers varied from two to eight weeks. The student groups whose teachers implemented the full range of LLV curriculum and assessment materials with students on a daily basis for at least 5 weeks had statistically significant learning gains across the three learner types: grade level readers, at-risk readers, and English language learners.

Characteristics of Treatment Group Participants

The majority of the treatment group students were grade level readers. The treatment groups also included 12% new English language learners and 24% at-risk readers whose native language was English. The One-sample Kolmogorov-Smirnov (1-Sample K-S) test, Kaiser-Meyer-Olkin Measure of Sampling Adequacy and the Bartlett's test of sphericity tests were conducted on all of the grade level groupings. These data showed that all participant groups fell within the normal population distribution and met the requirements for comparability of the groups. A few outliers were found with the pre-test data and these student cases were deleted from the data set analyzed for this report. Student cases were also deleted when there was no matched pair pre- and post-test due to student mobility. The sample population, although normal, did show a wide range and variability on pre-test performance. The frequency of each learner type for each grade level group is compiled in Appendix B.

Statistical Results of the Validation Study

An Independent T-Test was conducted on the validated pre- and post-tests for Grades 1, 2, 3, 4, and 5. Table 1 shows the statistical results for student groups who had opportunities to learn with the LLV Windows on Literacy (grades 1-3) and LLV Reading Expeditions (grades 4-5) curricula. These data exclude results from 1st Grade Treatment Groups 2 and 3, and 4th Grade Treatment Groups 1 and 2 because their teachers did not fully implement the LLV curriculum as design. Details about teachers' implementation of the LLV curriculum are described in the Classroom Adoption Practices section of this report. The Independent T-Test results comparing the mean difference in student gains for the Windows on Literacy (WOL) Treatment and Control Groups is $t(219) = 3.995$ and for Reading Expeditions (RE) Treatment and Control Groups is $t(154) = 4.084$; $p < 0.01$. These data indicate that students whose teachers fully implemented the LLV curriculum had statistically significant gains in their reading achievement compared to students in the control groups.

Table 1. Statistical findings compare the pre- and post-test differences for the WOL and RE treatment groups and their control groups.

	t	df	Independent T-Test for Equality of Means				95% Confidence Interval of the Difference	
			Sig. (1-tailed)	Mean Difference	Std. Error Difference	Lower	Upper	
WOL Treatment Group compared with WOL Control Group (equal variances assumed)	3.995*	219	.001	1.282	.321	.650	1.915	
RE Treatment Group compared with RE Control Group (equal variances assumed)	4.084*	154	.001	2.111	.517	1.090	3.132	

* indicates significant difference between pre- and post-test performance when $p < 0.01$

In Table 2, the statistical results include data from all 15 treatment classes and aggregated control group results. Significant gains in students' content area literacy development were found for 10 of 15 treatment groups as indicated by the asterisk in the “t” column. None of the control groups for each grade had significant learning gains as measured by the LLV criterion-referenced test.

The t-test results in Table 2 show a strong teacher effect. When teachers fully implemented the LLV curriculum as designed, students from all three learner type categories made progress in their reading achievement. Each grade level was assigned a set of 5 weekly units within the LLV curriculum. The content of all units pertained to science and social studies. Additional statistical tables comparing pre- and post-test means for each teacher group are compiled in Appendix C. Frequency of pre- and post-test differences and histograms comparing differences in students' pre- and post-test performance are compiled in Appendix D and E, respectively.

The LLV first grade texts used in this study correspond to Basal levels PP1, PP2 PP3-Primer and Primer. These treatment units, for the 1st graders, focused on vocabulary building and comprehension strategies for drawing inferences and asking questions about text pertaining to weather and seasons, families, force and motion, animal bodies and food from plants. The Paired Samples T-Test results for the 1st Grade Treatment Group 1 is $t(14) = 1.871$, for Treatment Group 2 is $t(14) 2.320$ and for Treatment Group 3 is $t(19) 2.392$; $p < 0.01$. The control group's Paired Samples T-Test result is similar at $t(45) = 1.806$; $p < 0.01$. None of the treatment groups in the 1st grade showed a statistically significant gain on the pre- and post-tests. However, the mean difference between students' pre- and post-test performance for all three treatment groups was more positive than for the control group students. Table 1 in Appendix D tabulates the frequency and range of gains for each group in Grade 1. Treatment Group 1 had the best results with 60% of its students gaining one or more points. Fifty-three percent of students in Treatment Group 2 and 50% of students in Treatment Group 3 gained one or more points while only 43% of students in the control group gained one or more points.

Table 2. Results of the Paired Samples T-Test measuring differences between pre- and post-test performances for participants in the *Language, Literacy & Vocabulary!* 2006-07 study indicate positive learning gains among 10 of 15 treatment groups. Results are reports in classroom groupings per teacher.

	1 st Grade Paired Sample Differences					T	df	Sig. (1-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
				Lower	Upper			
1st Grade Treatment Group 1 n = 15	1.000	2.070	.535	-.146	2.146	1.871	14	.041
1st Grade Treatment Group 2 n = 15	.667	1.113	.287	.050	1.283	2.320	14	.018
1st Grade Treatment Group 3 n = 20	.900	1.683	.376	.112	1.688	2.392	19	.013
1st Grade Control Group n = 46	.543	2.041	.301	-.063	1.150	1.806	45	.039

	2 nd Grade Paired Differences							
2nd Grade Treatment Group 1 n = 16	2.500	2.556	.639	1.138	3.862	3.912*	15	.001
2nd Grade Treatment Group 2 n = 16	1.250	1.844	.461	.267	2.233	2.712*	15	.080
2nd Grade Treatment Group 3 n = 18	1.222	2.016	.475	.220	2.225	2.572*	17	.010
2nd Grade Control Group n = 33	.788	1.933	.336	.103	1.473	2.342	32	.013
	3 rd Grade Paired Differences							
3rd Grade Treatment Group 1 n = 19	2.526	3.133	.719	1.016	4.037	3.514*	18	.001
3rd Grade Treatment Group 2 n = 19	3.316	2.849	.654	1.943	4.689	5.073*	18	.001
3rd Grade Control Group n = 39	.923	2.579	.413	.087	1.759	2.235	38	.015

* indicates significant difference between pre- and post-test performance when $p < 0.01$

	4 th Grade Paired Differences							
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		T	df	Sig. (1-tailed)
				Lower	Upper			
4th Grade Treatment Group 1 n = 17	1.353	4.471	1.084	-.946	3.652	1.248	16	.115
4th Grade Treatment Group 2 n = 13	2.077	3.639	1.009	-.122	4.276	2.058	12	.031
4th Grade Treatment Group 3 n = 18	3.500	2.550	.601	2.232	4.768	5.824*	17	.001
4th Grade Treatment Group 4 n = 17	3.588	3.337	.809	1.873	5.304	4.434*	16	.001
4th Grade Control Group n = 39	1.359	4.374	.700	-.059	2.777	1.940	38	.030
	5 th Grade Paired Differences							
5th Grade Treatment Group 1 n = 19	2.421	3.271	.750	.844	3.998	3.226*	18	.002
5th Grade Treatment	4.000	2.335	.674	2.516	5.484	5.933*	11	.001

Group 2 n = 12									
5th Grade Treatment Group 3 n = 19	2.947	2.223	.510	1.876	4.019	5.779*	18	.001	
5th Grade Control Group n = 32	.813	2.571	.454	-.114	1.739	1.788	31	.042	

* indicates significant difference between pre- and post-test performance when $p < 0.01$

The 2nd grade LLV units correspond to Basal levels for Grades 1 and 1-2. These treatment units focused on vocabulary building and comprehension strategies for drawing inferences and conclusions, and predicting and determining importance about text pertaining to space, where people live, plant life, producing goods, and historical events. All three of the treatment groups for 2nd grade showed a significant gain on the pre- and post-tests. The Paired Samples T-Test result for Treatment Group 1 is $t(15) = 3.912$, for Treatment Group 2 is $t(15) 2.712$ and for Treatment Group 3 is $t(17) 2.572$; $p < 0.01$. The control group's Paired Samples T-Test result is $t(32) = 2.342$; $p < 0.01$, which does not indicate a significant difference in students' pre- and post-test performance. The percent of 2nd grade students who had a learning gain measured by the pre- and post-tests ranged from 62% to 78% in the treatment groups compared with 45% in the control group. Table 2 in Appendix D tabulates the frequency and range of gains for each group in Grade 2. The majority of students in the treatment groups gained between 1 to 3 points while the majority of control group students had zero or negative gains.

The 3rd grade LLV texts correspond to Basal levels for Grades 2 and 2-3. These treatment units focused on vocabulary building and comprehension strategies for determining importance and asking questions about text including maps and content pertaining to prehistoric life, producing goods, U.S. geography, and the changing earth. The Paired Samples T-Test result for the 3rd grade Treatment Group 1 is $t(18) = 3.514$ and for Treatment Group 2 is $t(18) 5.073$; $p < 0.01$ which indicates that these students performed significantly better on the post-test than they did on the pre-test. The control group's Paired Samples T-Test result is $t(38) = 2.235$; $p < 0.01$, which indicates no significant difference in pre- and post-test performance. Between 68% and 79% of students in the treatment groups showed some gain while only 59% of students in the control group showed any learning progress. Table 3 in Appendix D tabulates the frequency and range of gains for each group in Grade 3. The majority of students in both of the 3rd Grade treatment groups gained 3 or more points while only 31% of students in the control group had a gain of 3 or more points on the criterion-referenced test.

The 4th grade LLV texts were intermediate levels 4 through 8. These treatment units focused on vocabulary building and comprehension strategies for determining importance and drawing inferences from text about U.S. history, earth science, life science and physical science. The Paired Samples T-Test result for the Treatment Group 1 is $t(16) = 1.248$, for Treatment Group 2 is $t(12) 2.058$ which are statistically non-significant. For Treatment Group 3 the paired sample t-test result is $t(17) 5.824$ and for Treatment Group 4 is $t(16) 4.434$; $p < 0.01$. These results indicates that students in Groups 3 and 4 performed significantly better on their post-test than they did on the pre-test. The control group's Paired Samples T-Test result is $t(38) = 1.940$; $p < 0.01$, which indicates no significant difference in pre- and post-test performance. Students in treatment groups 1 and 2 performed on a par with students in the control group. In these three groups, between 39%

and 49% of the students showed zero or negative gains on the post-test. In Treatment Groups 3 and 4, between 78% and 82% of students showed gains ranging from 1 to 10 points on the post-test with the majority gaining 3 or more points. Refer to Table 4 in Appendix D for the frequency and range of gains for each group in Grade 4.

The 5th grade LLV texts were intermediate level 4 through 8. These treatment units focused on vocabulary building and comprehension strategies for asking questions about and visualizing the meaning of text in U.S. history, earth science, life science and physical science. The Paired Samples T-Test for Treatment Group 1 is $t(18) = 3.226$, for Treatment Group 2 is $t(11) 5.933$, for Treatment Group 3 is $t(18) 5.779$; $p < 0.01$ which indicates that students in all three treatment groups performed significantly better on the post-test than they did on the pre-test. The control group's Paired Samples T-Test result is $t(31) = 1.788$; $p < 0.01$, which indicates no significant difference between pre- and post-test performance. Thirty-seven percent of students in Treatment Group 1 gained 3 or more points on the post-test. Some 83% of students in Treatment Groups 2 and 58% of students in Treatment Group 3 gained 3 or more points while only 25% in the control group gained more than 2 points. Refer to Table 5 in Appendix D for the frequency and range of gains for each group in Grade 5.

The Paired Samples T-Test data in Table 2 indicates a strong teacher effect, which is discussed in more detail under the Classroom Adoption section of this report. The majority of students in treatment groups whose teachers fully implemented the LLV curriculum units during 5 or more weeks of instruction showed a significant increase in their vocabulary building and comprehension strategy use as measured by the pre- and post-test. Meanwhile, the control groups' performance is less consistent with 40% to 50% of students showing zero or negative gains (see frequency tables in Appendix D). Teachers' classroom adoption practices are not the only factor impacting students' learning results, however.

Table 3 compares the mean difference between pre- and post-test performance for students grouped according to learner type (grade level reader, at-risk readers in English, or English language learner) in each grade level. Each student's post-test was coded with a learner type by their teacher. Teachers also verified the number of each type of learner on their surveys and during exit interviews. The results in Table 3 show that treatment group students in all three learner-type categories out-performed control students at every grade level except for the at-risk students in 1st grade. More robust statistical analysis of learner-types is not possible here because of the low participation rates of at-risk readers and English language learners. Nonetheless, grouping students according to learner type as in Table 3 provides insight into how the LLV curriculum positively impacts students with special needs as well as those reading at grade level.

Table 3. Descriptive statistics measure differences between pre- and post-test performances for participants in the *Language, Literacy & Vocabulary!* 2006-07 school year study. Results are reports by type of learner group per grade level.

	Mean	Std. Deviation	Std. Error Mean	Range	
				Minimum	Maximum
1st Grade Treatment					
Grade level readers n = 33	.55	1.175	.205	-2	3
At-risk in English n = 11	1.36	1.963	.592	-2	6
English language learners n = 6	1.67	2.805	1.145	-2	5
1st Grade Control					
Grade level readers n = 24	.21	1.382	.282	-2	4
At-risk in English n = 4	3.50	2.646	1.323	1	7
English language learners n = 18	.33	2.223	.524	-5	4
2nd Grade Treatment					
Grade level readers n = 27	1.04	1.786	.344	-3	6
At-risk in English n = 9	2.22	2.048	.683	-2	5
English language learners n = 14	2.43	2.738	.732	-2	7
2nd Grade Control					
Grade level readers n = 22	.55	1.535	.327	-3	3
At-risk in English n = 7	1.29	2.430	.918	-2	5
English language learners n = 4	1.25	3.202	1.601	-1	6
3rd Grade Treatment					
Grade level readers n = 26	3.19	2.728	.535	-1	9
At-risk in English n = 12	2.33	3.525	1.018	-3	9
English language learners n = 0	0	0	0	0	0
3rd Grade Control					
Grade level readers n = 25	1.16	2.340	.468	-4	4
At-risk in English n = 14	.50	3.006	.803	-5	5
English language learners n = 0	0	0	0	0	0
	Mean	Std. Deviation	Std. Error Mean	Student Performance Range	
				Minimum	Maximum
4th Grade Treatment					

Grade level readers n = 45	3.20	3.348	.499	-4	10
At-risk in English n = 13	1.39	4.253	1.180	-4	11
English language learners n = 7	1.71	3.592	1.358	-2	8
4th Grade Control					
4th Grade Control grade level readers n = 29	1.90	4.716	.876	-5	11
4th Grade Control at-risk in English n = 8	.00	2.726	.964	-3	4
English language learners n = 2	-1.00	4.243	3.000	-4	2
5th Grade Treatment					
Grade level readers n = 32	3.06	2.384	.415	0	9
At-risk in English n = 14	2.73	3.515	.907	0	12
English language learners n = 2	4.00	1.414	1.000	3	5
5th Grade Control					
Grade level readers n = 15	.33	2.160	.558	-3	4
At-risk in English n = 15	1.47	2.997	.774	-4	5
English language learners n = 2	-.50	.707	.500	-1	0

LLV Classroom Adoption Practices

A comparative analysis of teachers' weekly implementation logs, pre- and post-surveys, classroom observations and exit interview data indicates that teachers in the treatment groups who fully implemented the LLV curriculum as design had significant learning gains among students. Ten of the 15 treatment teachers reported completing all five LLV units with students. Students in these classrooms were engaged in daily LLV lessons. The teachers in these classrooms used the full range of LLV curriculum and assessment components for 5 to 8 weeks, consecutively. All 10 of these teachers reported differentiating their instruction with whole group, small group and individual student instruction as needed for each learner type. Their survey results and classroom observations indicate a strong focus on teaching vocabulary development and comprehension strategies while using the LLV curriculum.

The treatment teachers in Grades 1 and 4, who were less successful according to the student outcome data in Table 2, did not use the LLV curriculum consistently. For example, 1st grade teachers in Treatment Groups 2 and 3 only used the curriculum once or twice a week with students. The 1st grade teacher for Treatment Group 1 used the LLV curriculum with students 3 or four times weekly and her student had better gains than in these other 1st grade groups. In the 4th grade, the teacher for Treatment Group 1 reported only using the LLV curriculum one or two days a week. The teacher for the 4th grade Treatment Group 2 reported only 2 weeks of instruction using the LLV curriculum. In addition, her classroom experienced a 60% mobility rate during this study. These five teachers reported using the LLV curriculum sporadically and focused on teaching its content rather than comprehension strategies while the other 10 treatment teachers whose students had significant gains implemented LLV more thoroughly. Years of teaching experiences does not appear to have impacted how these teachers implemented the LLV curriculum with the possible exception of the 4th grade teacher for Treatment Group 1 who was in her second year of teaching. The other 4 teachers whose students did not have significant gains had 5 or more years of teaching experience.

Teachers who had a majority of grade level readers in their classrooms reported in their exit interviews that the LLV curriculum materials were too easy for their students. Yet their students had significant gains on the post-test. These teachers reported continuing daily instruction because the LLV curriculum engaged students in reading that was enjoyable. With grade level readers, they used the LLV curriculum as reinforcement for comprehension strategies and observed students transferring use of strategies to more challenging academic texts. Grade level readers were able to use the LLV curriculum at their independent reading level, which freed teachers to work in small groups or individually with at-risk readers and English language learners. Since learner engagement and enjoyment contribute to students' taking ownership of their literacy development, the appeal of the LLV curriculum appear to be a contributing factor in the success found among treatment groups.

Classroom observation and interview data indicates that control group students received a variety of curriculum and instruction. These ranged from whole group basal reading lessons in science and social studies to reading literature. Control group teachers focused on phonics, daily reading and vocabulary building with non-fiction basals in the first and second grades. In the upper grades, teachers used guided reading approaches with basal texts and focused on content learning more than comprehension strategies. Control group teachers reported that their texts did not support differentiating learning for various types of learners.

Conclusions and Recommendations

- The large learning gains found among the regular education students who were reading at grade level during the treatment show the value of providing students with content area curriculum materials that are enjoyable as well as informative. Grade level readers can easily use the LLV materials to independently study science and social studies content, which frees teachers to focus their attention on special needs students and to differentiate their instruction within a diverse group of students.
- It is recommended that a follow up study of the LLV curriculum be conducted with special needs students, i.e. at-risk readers in English and English language learners. The data about these types of learners, in this report, was limited due to high mobility of students and the early withdrawal from the study of one school with a high Hispanic student population. For purposes of statistical analysis, a larger sample size of special needs students is required to verify whether or not the positive impact LLV showed in this study holds true for the larger population of at-risk readers and English language learners.
- While the data indicates that the strong teacher effect in this study was due to differences in how thoroughly each teacher implemented the LLV curriculum as designed, it may also indicate a need for more professional development among teachers who are not skilled at differentiating instruction with a diversity of learner types in their classroom.

Appendix A: Test Validation

Validation of the Criterion-Reference Test

Statistical analyses were conducted using classical methods for establishing the validity and reliability of criterion-referenced tests. The purpose of the analysis was to determine the validity and reliability of the tests used to measure students' learning in the quasi-experimental design study of the *Language, Literacy and Vocabulary!* curriculum. The methods used to analyze the tests included the following statistical steps for each grade level:

- Determining whether or not the class groups have any outliers.
- Determining the normality of the sample population using the One-Sample Kolmogorov-Smirnov Test. This test was also used for each class group.
- Determining the suitability of the data for factor analysis using the Kaiser-Meyer-Olkin and the Bartlett's test of sphericity.
- Conducting a principal component factor analysis of the vocabulary and comprehension subscales using eigenvalues over 1 for factor selection and the Varimax rotation method with Kaiser Normalization.
- Conducting the Pearson correlation test to verify the strength of the correlation and reliability among the test items and total test score.
- Conducting the Cronbach Alpha test for reliability.
- Power analysis of each grade level test.
- Generating basic statistics such as analysis of means, standard deviation, effect size, and standard error of measurement.

All class groups across all grades were found to be within the normal range of distribution. A few outliers were found and these cases were deleted from the data set analyzed. All of the tests were found to be suitable for factor analysis. Review of the factor analysis rotated component matrices and Pearson correlation tables indicated a few test items needed to be eliminated to make the tests stronger measures of student learning. When a test item is found to be a weak measure it is removed from the final pre- and post-test statistical calculations for each grade level. Table 1 shows the Cronbach Alpha for each test and the test items used. The statistical analyses of the criterion-referenced test items indicated medium to high reliability for these tests. Table 2 shows the basic descriptive statistics for the validated set of test items. Additional statistics for each grade level provide more in-depth analysis of each grade level sample population and test item analysis according to the classical methods for establishing the validity and reliability described in the bulleted list above. Overall, the tests for each grade level had strong correlations. Items that lacked a strong correlation are deleted from the data set.

Table 1. Test reliability statistics for each of the grade level tests used for the pre-test data analysis.

N	Grade Level	Cronbach's Alpha* Vocabulary Scale	Cronbach's Alpha* Comprehension Scale	Test Items**
144	First Grade	.746	n/a	VQ 1-15
103	Second Grade	.685	.536	VQ 1, 3-12; CQ 1-3
89	Third Grade	.621	.616	VQ 1-13, 15; CQ 1-6
162	Fourth Grade	.654	.528	VQ 1-14; CQ 1-6
122	Fifth Grade	.600	.547	VQ 1, 3-15; CQ 1-6

*> .5 indicates mid-level reliability; >.75 high-level reliability (Hinton, 2004)

** VQ = vocabulary questions; CQ = comprehension questions

Table 2. Power of the test* for the Paired Samples T-Test (1-tailed; post-hoc using pre-test data).

Grade Level	Power of Test	Effect Size	Alpha
1st Grade	0.636	.41	0.05
2nd Grade	0.999	1.13	0.05
3rd Grade	0.060	.02	0.05
4th Grade	0.911	.61	0.05
5th Grade	0.999	1.11	0.05

*Computed using the G-Power calculator

Reliability and Validity Analysis Grade 1

1st Grade Whole Group One-Sample Kolmogorov-Smirnov Test of Normality

		TotalScore
N		144
Normal Parameters(a,b)	Mean	12.90
	Std. Deviation	2.251
Most Extreme Differences	Absolute	.239
	Positive	.176
	Negative	-.239
Kolmogorov-Smirnov Z		2.873
Asymp. Sig. (2-tailed)		.000

a Test distribution is Normal. b Calculated from pre-test data.

1st Grade One-Sample Kolmogorov-Smirnov Test of Normality per Teacher Group

TotalScore		Control Group	Group 1	Group 2	Group 3
Normal Parameters(a,b)	Mean	83	17	18	26
	Std. Deviation	12.82	11.94	13.89	13.12
Most Extreme Differences	Absolute	2.343	2.883	1.023	1.883
	Positive	.254	.214	.210	.219
	Negative	.176	.144	.141	.158
Kolmogorov-Smirnov Z		-.254	-.214	-.210	-.219
Asymp. Sig. (2-tailed)		2.311	.882	.891	1.118
		.000	.417	.406	.164

a Test distribution is Normal. b Calculated from pre-test data.

1st grade Kaiser-Meyer-Olkin and the Bartlett's test of sphericity test results

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.702
Bartlett's Test of Sphericity	Approx. Chi-Square	702.351
	df	66
	Sig.	.000

1st Grade Factor Analysis for test items and total score show total variance explained by four factor.

	Component			
	1	2	3	4
TotalScore	.990	.069	.025	.005
VQ1	.352	.380	.010	-.405
VQ2	.638	-.098	.132	.117
VQ3	.572	-.054	.030	-.372
VQ4	.710	-.006	-.023	-.224
VQ5	.627	-.286	-.204	.132
VQ6	.630	.197	-.278	-.178
VQ7	.546	.077	.091	.392
VQ8	.626	.129	.322	-.046
VQ9	.411	-.237	.179	.531
VQ10	.485	-.319	-.585	-.094
VQ11	.297	.716	-.036	.216
VQ12	.565	.277	-.159	.222
VQ13	.627	-.100	.199	.086
VQ14	.576	-.244	-.178	-.046
VQ15	.374	-.232	.610	-.298

Extraction Method: Principal Component Analysis.

1st Pearson Correlations among individual test items and total test shows high reliability among the test items and total test score

n = 146		Total Pre-Test Score
VQ1	Pearson Correlation	.388(**)
	Sig. (2-tailed)	.000
VQ2	Pearson Correlation	.593(**)
	Sig. (2-tailed)	.000
VQ4	Pearson Correlation	.682(**)
	Sig. (2-tailed)	.000
VQ5	Pearson Correlation	.577(**)
	Sig. (2-tailed)	.000
VQ6	Pearson Correlation	.619(**)
	Sig. (2-tailed)	.000
VQ7	Pearson Correlation	.549(**)
	Sig. (2-tailed)	.000
VQ8	Pearson Correlation	.593(**)
	Sig. (2-tailed)	.000
VQ12	Pearson Correlation	.569(**)
	Sig. (2-tailed)	.000
VQ13	Pearson Correlation	.598(**)
	Sig. (2-tailed)	.000
VQ14	Pearson Correlation	.587(**)
	Sig. (2-tailed)	.000
VQ15	Pearson Correlation	.416(**)
	Sig. (2-tailed)	.000

** Correlation is significant at the 0.01 level (2-tailed).
VQ = vocabulary question

Reliability and Validity Analysis Grade 2 / Tests for Normality of the 2nd Grade Groups

2nd Grade Whole Group One-Sample Kolmogorov-Smirnov Test of Normality

		TotalScore
N		103
Normal Parameters(a,b)	Mean	11.05
	Std. Deviation	2.795
Most Extreme Differences	Absolute	.134
	Positive	.087
	Negative	-.134
Kolmogorov-Smirnov Z		1.358
Asymp. Sig. (2-tailed)		.050

a Test distribution is Normal. b Calculated from pre-test data.

2nd Grade One-Sample Kolmogorov-Smirnov Test of Normality per Teacher Group

TotalScore		Control Group	Group 1	Group 2	Group 3
		42	19	17	25
Normal Parameters(a,b)	Mean	11.31	8.16	12.24	12.00
	Std. Deviation	2.384	3.354	1.921	1.979
Most Extreme Differences	Absolute	.162	.113	.137	.213
	Positive	.119	.093	.137	.156
	Negative	-.162	-.113	-.115	-.213
Kolmogorov-Smirnov Z		1.047	.492	.565	1.067
Asymp. Sig. (2-tailed)		.223	.969	.907	.205

a Test distribution is Normal. b Calculated from pre-test data.

2nd Grade Kaiser-Meyer-Olkin and the Bartlett's test of sphericity vocabulary test item results.

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.539
Bartlett's Test of Sphericity	Approx. Chi-Square	428.856
	df	78
	Sig.	.000

2nd Grade Kaiser-Meyer-Olkin and the Bartlett's test of sphericity comprehension test item results.

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.669
Bartlett's Test of Sphericity	Approx. Chi-Square	105.084
	df	6
	Sig.	.000

2nd Grade Factor Analysis for the 12 vocabulary test items and total score.

	Component			
	1	2	3	4
TotalScore	.713	.527	.261	.312
VQ1	.026	.687	.004	-.087
VQ2	-.026	-.051	.806	.031
VQ3	.072	.265	.525	.034
VQ4	.220	-.047	.657	.468
VQ5	-.092	.083	.154	.787
VQ6	.109	.454	.258	.504
VQ7	.153	.645	.118	.122
VQ8	.536	.303	.036	.308
VQ9	.808	-.076	.219	-.157
VQ10	.475	-.108	-.377	.533
VQ11	.335	.475	-.039	.175
VQ12	.714	.305	-.082	-.027

Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization.

2nd Grade Factor Analysis for 3 comprehension test items and total score show total variance explained by one factor.

	Component
	1
TotalScore	.887
CQ1	.780
CQ2	.706
CQ3	.607

Extraction Method: Principal Component Analysis. a 1 components extracted.

2nd Grade Pearson Correlations among individual test items and total test score shows high reliability among 14 of the 15 test items and total test score

N = 103		TotalScore
VQ1	Pearson Correlation	.354(**)
	Sig. (2-tailed)	.000
VQ2	Pearson Correlation	.198
	Sig. (2-tailed)	.045
VQ3	Pearson Correlation	.312(**)
	Sig. (2-tailed)	.001
VQ4	Pearson Correlation	.422(**)
	Sig. (2-tailed)	.000
VQ5	Pearson Correlation	.275(**)
	Sig. (2-tailed)	.005
VQ6	Pearson Correlation	.525(**)
	Sig. (2-tailed)	.000
VQ7	Pearson Correlation	.525(**)
	Sig. (2-tailed)	.000
VQ8	Pearson Correlation	.635(**)
	Sig. (2-tailed)	.000
VQ9	Pearson Correlation	.547(**)

N = 103		TotalScore
	Sig. (2-tailed)	.000
VQ10	Pearson Correlation	.353(**)
	Sig. (2-tailed)	.000
VQ11	Pearson Correlation	.528(**)
	Sig. (2-tailed)	.000
VQ12	Pearson Correlation	.612(**)
	Sig. (2-tailed)	.000

** Correlation is significant at the 0.01 level (2-tailed).
VQ = vocabulary question;

2nd Grade Pearson Correlations among the comprehension test items and total test score

n = 103		TotalScore
CQ1	Pearson Correlation	.571(**)
	Sig. (2-tailed)	.000
CQ2	Pearson Correlation	.557(**)
	Sig. (2-tailed)	.000
CQ3	Pearson Correlation	.462(**)
	Sig. (2-tailed)	.000

** Correlation is significant at the 0.01 level (2-tailed).
CQ = comprehension question

Reliability and Validity Analysis Grade 3

3rd Grade Whole Group One-Sample Kolmogorov-Smirnov Test of Normality

		TotalScore
N		89
Normal Parameters(a,b)	Mean	12.94
	Std. Deviation	3.773
Most Extreme Differences	Absolute	.083
	Positive	.056
	Negative	-.083
Kolmogorov-Smirnov Z		.785
Asymp. Sig. (2-tailed)		.569

a Test distribution is Normal. b Calculated from pre-test data.

3rd Grade One-Sample Kolmogorov-Smirnov Test of Normality per Teacher Group

TotalScore		Control Group	Group 1	Group 2
N =		45	24	20
Normal Parameters(a,b)	Mean	13.20	11.96	13.55
	Std. Deviation	3.609	4.288	3.426
Most Extreme Differences	Absolute	.114	.141	.125
	Positive	.084	.126	.125
	Negative	-.114	-.141	-.113
Kolmogorov-Smirnov Z		.767	.692	.557
Asymp. Sig. (2-tailed)		.598	.724	.916

a Test distribution is Normal. b Calculated from pre-test data.

3rd Grade Kaiser-Meyer-Olkin and the Bartlett's test of sphericity test results for the vocabulary scale

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.533
Bartlett's Test of Sphericity	Approx. Chi-Square	344.660
	df	120
	Sig.	.000

3rd Grade Kaiser-Meyer-Olkin and Bartlett's test of sphericity test results for comprehension scale

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.698
Bartlett's Test of Sphericity	Approx. Chi-Square	163.560
	df	21
	Sig.	.000

3rd Grade Factor Analysis for the vocabulary test items and total test score.

	Component					
	1	2	3	4	5	6
TotalScore	.786	.362	.128	.261	.250	.146
VQ1	.603	-.089	.101	-.036	.228	.089
VQ2	.623	.168	.094	-.294	.100	.041
VQ3	.181	-.096	-.479	.225	-.219	.585
VQ4	.618	.035	-.006	.186	-.200	-.059
VQ5	.255	.416	.288	.464	-.096	-.134
VQ6	.438	.364	.308	.112	.263	.015
VQ7	.335	.632	.124	-.038	.055	-.232
VQ8	.102	.039	.233	-.140	.004	.814
VQ9	-.086	-.035	-.108	.404	.701	.104
VQ10	.745	.016	-.011	.149	-.040	.066
VQ11	.307	-.268	.648	.155	-.199	-.101
VQ12	.102	.008	.027	.784	.109	-.021
VQ13	.251	.095	.102	-.133	.693	-.197
VQ14	-.005	-.165	-.740	-.007	-.094	-.216
VQ15	-.082	.826	-.090	.031	.008	.149

Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization.

3rd Grade Factor Analysis for the comprehension and total test score.

	Component		
	1	2	3
TotalScore	.695	.522	.300
CQ1	-.009	.787	.220
CQ2	.057	.062	.933
CQ3	.715	.349	-.146
CQ4	.600	.182	.450
CQ5	.211	.717	-.077
CQ6	.826	-.112	.072

Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization.

3rd Grade Pearson Correlations among the vocabulary test items and total test score

n = 89		TotalScore
VQ1	Pearson Correlation	.484(**)
	Sig. (2-tailed)	.000
VQ2	Pearson Correlation	.481(**)
	Sig. (2-tailed)	.000
VQ3	Pearson Correlation	.142
	Sig. (2-tailed)	.185
VQ4	Pearson Correlation	.457(**)
	Sig. (2-tailed)	.000
VQ5	Pearson Correlation	.485(**)
	Sig. (2-tailed)	.000
VQ6	Pearson Correlation	.586(**)
	Sig. (2-tailed)	.000
VQ7	Pearson Correlation	.512(**)
	Sig. (2-tailed)	.000
VQ8	Pearson Correlation	.240
	Sig. (2-tailed)	.023
VQ9	Pearson Correlation	.206
	Sig. (2-tailed)	.053
VQ10	Pearson Correlation	.585(**)
	Sig. (2-tailed)	.000
VQ11	Pearson Correlation	.261
	Sig. (2-tailed)	.014
VQ12	Pearson Correlation	.276(**)
	Sig. (2-tailed)	.009
VQ13	Pearson Correlation	.368(**)
	Sig. (2-tailed)	.000
VQ14	Pearson Correlation	-.147
	Sig. (2-tailed)	.168
VQ15	Pearson Correlation	.231
	Sig. (2-tailed)	.029

** Correlation is significant at the 0.01 level (2-tailed).

3rd Grade Pearson Correlations among the comprehension test items and total test score

n = 89		TotalScore
CQ1	Pearson Correlation	.438(**)
	Sig. (2-tailed)	.000
CQ2	Pearson Correlation	.329(**)
	Sig. (2-tailed)	.002
CQ3	Pearson Correlation	.625(**)
	Sig. (2-tailed)	.000
CQ4	Pearson Correlation	.602(**)
	Sig. (2-tailed)	.000
CQ5	Pearson Correlation	.427(**)
	Sig. (2-tailed)	.000
CQ6	Pearson Correlation	.466(**)
	Sig. (2-tailed)	.000

** Correlation is significant at the 0.01 level (2-tailed).

Reliability and Validity Analysis Grade 4

4th Grade Whole Group One-Sample Kolmogorov-Smirnov Test of Normality

		TotalScore
N		162
Normal Parameters(a,b)	Mean	14.59
	Std. Deviation	3.513
Most Extreme Differences	Absolute	.125
	Positive	.068
	Negative	-.125
Kolmogorov-Smirnov Z		1.588
Asymp. Sig. (2-tailed)		.013

a Test distribution is Normal. b Calculated from pre-test data.

4th Grade One-Sample Kolmogorov-Smirnov Test of Normality per Teacher Group

TotalScore		Control Group	Group 1	Group 2	Group 3	Group 4	Group 5
N =		46	24	22	22	21	27
Normal Parameters(a,b)	Mean	15.50	12.25	12.68	13.14	15.43	17.22
	Std. Deviation	3.031	3.881	3.772	2.455	3.124	2.044
Most Extreme Differences	Absolute	.121	.177	.140	.160	.144	.210
	Positive	.093	.177	.120	.090	.110	.210
	Negative	-.121	-.095	-.140	-.160	-.144	-.123
Kolmogorov-Smirnov Z		.823	.868	.655	.749	.660	1.091
Asymp. Sig. (2-tailed)		.507	.438	.784	.629	.777	.185

a Test distribution is Normal. b Calculated from pre-test data.

4th Grade Kaiser-Meyer-Olkin and the Bartlett's test of sphericity results for the vocabulary subscale

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.496
Bartlett's Test of Sphericity	Approx. Chi-Square	618.944
	df	120
	Sig.	.000

4th Grade Kaiser-Meyer-Olkin and the Bartlett's test of sphericity results for the comprehension subscale

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.697
Bartlett's Test of Sphericity	Approx. Chi-Square	208.732
	df	21
	Sig.	.000

4th Grade Factor Analysis for the vocabulary subscale and total score

	Component				
	1	2	3	4	5
TotalScore	.528	.400	.437	.428	.357
VQ1	-.041	.020	.800	.097	.050
VQ2	-.052	.782	-.001	-.045	-.087
VQ3	.274	.485	.279	-.250	.002
VQ4	.096	-.026	.352	.488	.247
VQ5	.272	-.037	.267	.689	.091
VQ6	.600	.256	.093	.127	-.119
VQ7	.431	.147	.080	.105	.170
VQ8	.585	-.266	.243	-.147	.325
VQ9	.384	.485	.091	.130	.017
VQ10	.172	.558	-.053	.108	.375
VQ11	.045	.041	-.134	.765	-.118
VQ12	.715	.015	-.085	.135	-.123
VQ13	-.066	.305	.149	.066	.672
VQ14	.165	.081	.665	.033	-.106
VQ15	.046	-.177	-.154	-.014	.728

Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization.

4th Grade Factor Analysis for the comprehension subscale and total score

	Component	
	1	2
TotalScore	.731	.492
CQ1	.555	.298
CQ2	-.155	.850
CQ3	.537	.356
CQ4	.373	.469
CQ5	.743	-.151
CQ6	.638	-.026

Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization.

4th Grade Pearson Correlations among the vocabulary test items and total test score

n = 162		TotalScore
VQ1	Pearson Correlation	.388(**)
	Sig. (2-tailed)	.000
VQ2	Pearson Correlation	.239(**)
	Sig. (2-tailed)	.002
VQ3	Pearson Correlation	.322(**)
	Sig. (2-tailed)	.000
VQ4	Pearson Correlation	.449(**)
	Sig. (2-tailed)	.000
VQ5	Pearson Correlation	.540(**)
	Sig. (2-tailed)	.000
VQ6	Pearson Correlation	.448(**)
	Sig. (2-tailed)	.000
VQ7	Pearson Correlation	.434(**)
	Sig. (2-tailed)	.000
VQ8	Pearson Correlation	.373(**)
	Sig. (2-tailed)	.000
VQ9	Pearson Correlation	.488(**)
	Sig. (2-tailed)	.000
VQ10	Pearson Correlation	.465(**)
	Sig. (2-tailed)	.000
VQ11	Pearson Correlation	.289(**)
	Sig. (2-tailed)	.000
VQ12	Pearson Correlation	.355(**)
	Sig. (2-tailed)	.000
VQ13	Pearson Correlation	.417(**)
	Sig. (2-tailed)	.000
VQ14	Pearson Correlation	.407(**)
	Sig. (2-tailed)	.000
VQ15	Pearson Correlation	.130
	Sig. (2-tailed)	.098

** Correlation is significant at the 0.01 level (2-tailed).

4th Grade Pearson Correlations among the comprehension test items and total test score

		TotalScore
CQ1	Pearson Correlation	.418(**)
	Sig. (2-tailed)	.000
CQ2	Pearson Correlation	.283(**)
	Sig. (2-tailed)	.000
CQ3	Pearson Correlation	.527(**)
	Sig. (2-tailed)	.000
CQ4	Pearson Correlation	.414(**)
	Sig. (2-tailed)	.000
CQ5	Pearson Correlation	.444(**)
	Sig. (2-tailed)	.000
CQ6	Pearson Correlation	.422(**)
	Sig. (2-tailed)	.000

** Correlation is significant at the 0.01 level (2-tailed).

Reliability and Validity Analysis Grade 5

5th Grade Whole Group One-Sample Kolmogorov-Smirnov Test of Normality

		TotalScore
N		122
Normal Parameters(a,b)	Mean	12.42
	Std. Deviation	3.604
Most Extreme Differences	Absolute	.113
	Positive	.068
	Negative	-.113
Kolmogorov-Smirnov Z		1.252
Asymp. Sig. (2-tailed)		.087

a Test distribution is Normal. b Calculated from pre-test data.

5th Grade One-Sample Kolmogorov-Smirnov Test of Normality per Teacher Group

TotalScore		Control Group	Group 1	Group 2	Group 3	Group 4
n =		37	22	22	21	20
Normal Parameters(a,b)	Mean	13.30	10.32	11.27	12.48	14.30
	Std. Deviation	2.788	4.110	4.119	3.683	2.105
Most Extreme Differences	Absolute	.130	.133	.163	.184	.182
	Positive	.130	.081	.123	.113	.182
	Negative	-.070	-.133	-.163	-.184	-.100
Kolmogorov-Smirnov Z		.792	.624	.762	.845	.812
Asymp. Sig. (2-tailed)		.557	.831	.607	.474	.525

a Test distribution is Normal. b Calculated from pre-test data.

5th Grade Kaiser-Meyer-Olkin and the Bartlett's test of sphericity test results for the vocabulary subscale

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.553
Bartlett's Test of Sphericity	Approx. Chi-Square	427.165
	df	120
	Sig.	.000

5th Grade Kaiser-Meyer-Olkin and the Bartlett's test of sphericity test results for the comprehension subscale

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.692
Bartlett's Test of Sphericity	Approx. Chi-Square	143.110
	df	21
	Sig.	.000

5th Grade Factor Analysis for the vocabulary subscale and total score

	Component					
	1	2	3	4	5	6
TotalScore	.637	.595	.287	.084	.271	.049
VQ1	.001	.018	.031	.186	.849	-.007
VQ2	-.042	-.091	.677	-.156	.369	-.010
VQ3	.072	.153	.718	.117	-.195	.155
VQ4	.399	.148	-.030	-.647	-.114	-.091
VQ5	.800	-.163	.097	-.067	.017	-.068
VQ6	.684	.361	-.035	.173	-.084	.067
VQ7	-.043	.781	-.062	-.017	-.040	-.024
VQ8	.465	.520	-.033	.024	-.053	-.264
VQ9	.440	.180	-.317	.048	.286	.393
VQ10	.133	.620	.281	.324	.014	.138
VQ11	.096	.483	.021	-.374	.346	.018
VQ12	.058	.109	-.050	.496	.058	-.085
VQ13	.342	.105	.408	.035	.067	-.424
VQ14	.393	.042	.081	.592	.031	.002
VQ15	.005	.002	.168	-.054	-.011	.846

Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization.

5th Grade Factor Analysis for the comprehension subscale and total score; only one factor was found.

Component	Initial Eigenvalues		
	Total	% of Variance	Cumulative %
1	2.521	36.010	36.010
2	.993	14.188	50.198
3	.908	12.978	63.176
4	.865	12.360	75.536
5	.805	11.496	87.031
6	.590	8.429	95.461
7	.318	4.539	100.000

Extraction Method: Principal Component Analysis.

5th Grade Pearson Correlations among the vocabulary test items and total test score

n = 122		TotalScore
VQ1	Pearson Correlation	.251(**)
	Sig. (2-tailed)	.005
VQ2	Pearson Correlation	.195
	Sig. (2-tailed)	.031
VQ3	Pearson Correlation	.273(**)
	Sig. (2-tailed)	.002
VQ4	Pearson Correlation	.255(**)
	Sig. (2-tailed)	.005
VQ5	Pearson Correlation	.436(**)
	Sig. (2-tailed)	.000
VQ6	Pearson Correlation	.599(**)
	Sig. (2-tailed)	.000
VQ7	Pearson Correlation	.406(**)
	Sig. (2-tailed)	.000
VQ8	Pearson Correlation	.558(**)
	Sig. (2-tailed)	.000
VQ9	Pearson Correlation	.359(**)
	Sig. (2-tailed)	.000
VQ10	Pearson Correlation	.559(**)
	Sig. (2-tailed)	.000
VQ11	Pearson Correlation	.403(**)
	Sig. (2-tailed)	.000
VQ12	Pearson Correlation	.159
	Sig. (2-tailed)	.080
VQ13	Pearson Correlation	.385(**)
	Sig. (2-tailed)	.000
VQ14	Pearson Correlation	.369(**)
	Sig. (2-tailed)	.000
VQ15	Pearson Correlation	.104
	Sig. (2-tailed)	.254

** Correlation is significant at the 0.01 level (2-tailed).

5th Grade Pearson Correlations among the comprehension test items and total test score

n = 122		TotalScore
CQ1	Pearson Correlation	.414(**)
	Sig. (2-tailed)	.000
CQ2	Pearson Correlation	.426(**)
	Sig. (2-tailed)	.000
CQ3	Pearson Correlation	.360(**)
	Sig. (2-tailed)	.000
CQ4	Pearson Correlation	.434(**)
	Sig. (2-tailed)	.000
CQ5	Pearson Correlation	.388(**)
	Sig. (2-tailed)	.000
CQ6	Pearson Correlation	.448(**)
	Sig. (2-tailed)	.000

** Correlation is significant at the 0.01 level (2-tailed).

Appendix B: Descriptive Statistics by Learner Type

Table 1. Frequency and percent of learner type in 1st grade

1st Grade Treatment Group 1	Frequency	Percent	Valid Percent	Cumulative Percent
Grade level reader	4	26.7	26.7	26.7
At-risk in English	5	33.3	33.3	60.0
ELL	6	40.0	40.0	100.0
Total	15	100.0	100.0	
1st Grade Treatment Group 2	Frequency	Percent	Valid Percent	Cumulative Percent
Grade level reader	15	100.0	100.0	100.0
1st Grade Treatment Group 3	Frequency	Percent	Valid Percent	Cumulative Percent
Grade level reader	14	70.0	70.0	70.0
At-risk in English	6	30.0	30.0	100.0
Total	20	100.0	100.0	
1st Grade Control Group	Frequency	Percent	Valid Percent	Cumulative Percent
Grade level reader	24	52.2	52.2	52.2
At-risk in English	4	8.7	8.7	60.9
ELL	18	39.1	39.1	100.0
Total	46	100.0	100.0	

Table 2. Frequency and percent of learner type in 2nd grade

2nd Grade Treatment Group 1	Frequency	Percent	Valid Percent	Cumulative Percent
At-risk in English	2	12.5	12.5	12.5
ELL	14	87.5	87.5	100.0
Total	16	100.0	100.0	
2nd Grade Treatment Group 2	Frequency	Percent	Valid Percent	Cumulative Percent
Grade level reader	12	75.0	75.0	75.0
At-risk in English	4	25.0	25.0	100.0
Total	16	100.0	100.0	
2nd Grade Treatment Group 3	Frequency	Percent	Valid Percent	Cumulative Percent
Grade level reader	15	83.3	83.3	83.3
At-risk in English	3	16.7	16.7	100.0
Total	18	100.0	100.0	
2nd Grade Control Group	Frequency	Percent	Valid Percent	Cumulative Percent
Grade level reader	22	66.7	66.7	66.7
At-risk in English	7	21.2	21.2	87.9
ELL	4	12.1	12.1	100.0
Total	33	100.0	100.0	

Table 3. Frequency and percent of learner type in 3rd grade

3rd Grade Treatment Group 1	Frequency	Percent	Valid Percent	Cumulative Percent
Grade level reader	10	52.6	52.6	52.6
At-risk in English	9	47.4	47.4	100.0
Total	19	100.0	100.0	
3rd Grade Treatment Group 2	Frequency	Percent	Valid Percent	Cumulative Percent
Grade level reader	16	84.2	84.2	84.2
At-risk in English	3	15.8	15.8	100.0
Total	19	100.0	100.0	
3rd Grade Control Group	Frequency	Percent	Valid Percent	Cumulative Percent
Grade level reader	25	64.1	64.1	64.1
At-risk in English	14	35.9	35.9	100.0
Total	39	100.0	100.0	

Table 4. Frequency and percent of learner type in 4th grade

4th Grade Treatment Group 1	Frequency	Percent	Valid Percent	Cumulative Percent
Grade level reader	4	23.5	23.5	23.5
At-risk in English	12	70.6	70.6	94.1
ELL	1	5.9	5.9	100.0
Total	17	100.0	100.0	
4th Grade Treatment Group 2	Frequency	Percent	Valid Percent	Cumulative Percent
Grade level reader	9	69.2	69.2	69.2
ELL	4	30.8	30.8	100.0
Total	13	100.0	100.0	
4th Grade Treatment Group 3	Frequency	Percent	Valid Percent	Cumulative Percent
Grade level reader	16	88.9	88.9	88.9
ELL	2	11.1	11.1	100.0
Total	18	100.0	100.0	
4th Grade Treatment Group 4	Frequency	Percent	Valid Percent	Cumulative Percent
Grade level reader	16	94.1	94.1	94.1
At-risk in English	1	5.9	5.9	100.0
Total	17	100.0	100.0	
4th Grade Control Group	Frequency	Percent	Valid Percent	Cumulative Percent
Grade level reader	29	74.4	74.4	74.4
At-risk in English	8	20.5	20.5	94.9
ELL	2	5.1	5.1	100.0
Total	39	100.0	100.0	

Table 5. Frequency and percent of learner type in 5th grade

5th Grade Treatment Group 1	Frequency	Percent	Valid Percent	Cumulative Percent
Grade level reader	3	15.8	15.8	15.8
At-risk in English	14	73.7	73.7	89.5
ELL	2	10.5	10.5	100.0
Total	19	100.0	100.0	
5th Grade Treatment Group 2	Frequency	Percent	Valid Percent	Cumulative Percent
Grade level reader	11	91.7	91.7	91.7
At-risk in English	1	8.3	8.3	100.0
Total	12	100.0	100.0	
5th Grade Treatment Group 3	Frequency	Percent	Valid Percent	Cumulative Percent
Grade level reader	19	100.0	100.0	100.0
5th Grade Control Group	Frequency	Percent	Valid Percent	Cumulative Percent
Grade level reader	15	46.9	46.9	46.9
At-risk in English	15	46.9	46.9	93.8
ELL	2	6.3	6.3	100.0
Total	32	100.0	100.0	

Appendix C: Descriptive Statistics for Grade Level Pre- and Post-Test Performance

1st Grade Descriptive Statistics						
	N	Minimum	Maximum	Mean	Std. Deviation	Variance
Treatment Group 1						
Pre-Test	15	6	15	12.27	2.604	6.781
Post-Test		11	15	13.27	1.335	1.781
Treatment Group 2						
Pre-Test	15	12	15	13.93	1.033	1.067
Post-Test		14	15	14.60	.507	.257
Treatment Group 3						
Pre-Test	20	8	15	13.55	1.638	2.682
Post-Test		13	15	14.45	.686	.471
Control Group						
Pre-Test	46	4	15	12.87	2.535	6.427
Post-Test		8	15	13.41	1.857	3.448
2nd Grade Descriptive Statistics						
	N	Minimum	Maximum	Mean	Std. Deviation	Variance
Treatment Group 1						
Pre-Test	16	2	13	6.69	3.156	9.963
Post-Test		5	13	9.19	2.639	6.963
Treatment Group 2						
Pre-Test	16	8	14	11.38	1.893	3.583
Post-Test		8	14	12.63	1.708	2.917
Treatment Group 3						
Pre-Test	18	6	13	10.61	2.090	4.369
Post-Test		9	14	11.83	1.425	2.029
Control Group						
Pre-Test	33	7	14	10.73	2.212	4.892
Post-Test		6	14	11.52	1.955	3.820
3rd Grade Descriptive Statistics						
	N	Minimum	Maximum	Mean	Std. Deviation	Variance
Treatment Group 1						
Pre-Test	19	5	19	13.00	3.651	13.333
Post-Test		9	20	15.53	3.454	11.930
Treatment Group 2						
Pre-Test	19	8	20	13.21	3.552	12.620
Post-Test		12	20	16.53	2.318	5.374
Control Group						
Pre-Test	39	5	19	13.13	3.614	13.062
Post-Test		6	19	14.05	3.402	11.576

4th Grade Descriptive Statistics						
	N	Minimum	Maximum	Mean	Std. Deviation	Variance
Treatment Group 1						
Pre-Test	17	7	20	13.65	4.271	18.243
Post-Test		5	18	15.00	3.428	11.750
Treatment Group 2						
Pre-Test	13	8	18	14.00	3.317	11.000
Post-Test		12	20	16.08	2.362	5.577
Treatment Group 3						
Pre-Test	18	10	18	14.83	2.256	5.088
Post-Test		16	20	18.33	1.237	1.529
Treatment Group 4						
Pre-Test	17	7	17	13.41	3.183	10.132
Post-Test		13	20	17.00	1.936	3.750
Control Group						
Pre-Test	39	6	20	14.59	3.370	11.354
Post-Test		11	20	15.95	2.502	6.260
Fifth Grade Descriptive Statistics						
	N	Minimum	Maximum	Mean	Std. Deviation	Variance
Treatment Group 1						
Pre-Test	19	3	17	9.47	4.005	16.041
Post-Test		8	17	11.89	2.961	8.766
Treatment Group 2						
Pre-Test	12	7	18	13.08	3.397	11.538
Post-Test		12	20	17.08	2.575	6.629
Treatment Group 3						
Pre-Test	19	11	17	13.89	2.079	4.322
Post-Test		12	20	16.84	2.340	5.474
Control Group						
Pre-Test	32	8	19	13.13	2.814	7.919
Post-Test		8	19	13.94	2.435	5.931

Appendix D: Group Frequency Charts for Each Grade Level

Table 1. Frequency of pre- and post-test differences among 1st graders

Treatment Group 1	Frequency	Percent	Cumulative Percent
Valid -2	2	13.3	13.3
-1	2	13.3	26.7
0	2	13.3	40.0
1	3	20.0	60.0
2	3	20.0	80.0
3	1	6.7	86.7
4	1	6.7	93.3
5	1	6.7	100.0
Total	15	100.0	
Treatment Group 2	Frequency	Percent	Cumulative Percent
Valid -1	2	13.3	13.3
0	5	33.3	46.7
1	5	33.3	80.0
2	2	13.3	93.3
3	1	6.7	100.0
Total	15	100.0	
Treatment Group 3	Frequency	Percent	Cumulative Percent
Valid -2	1	5.0	5.0
-1	1	5.0	10.0
0	8	40.0	50.0
1	4	20.0	70.0
2	4	20.0	90.0
3	1	5.0	95.0
6	1	5.0	100.0
Total	20	100.0	
Control Group	Frequency	Percent	Cumulative Percent
Valid -5	1	2.2	2.2
-3	1	2.2	4.3
-2	3	6.5	10.9
-1	6	13.0	23.9
0	15	32.6	56.5
1	9	19.6	76.1
2	4	8.7	84.8
3	3	6.5	91.3
4	3	6.5	97.8
7	1	2.2	100.0
Total	46	100.0	

Table 2. Frequency of pre- and post-test differences among 2nd graders

Treatment Group 1	Frequency	Percent	Cumulative Percent
Valid -2	1	6.3	6.3
-1	1	6.3	12.5
0	2	12.5	25.0
1	2	12.5	37.5
2	1	6.3	43.8
3	4	25.0	68.8
4	1	6.3	75.0
5	2	12.5	87.5
6	1	6.3	93.8
7	1	6.3	100.0
Total	16	100.0	
Treatment Group 2	Frequency	Percent	Cumulative Percent
Valid -2	1	6.3	6.3
0	5	31.3	37.5
1	4	25.0	62.5
2	4	25.0	87.5
4	1	6.3	93.8
6	1	6.3	100.0
Total	16	100.0	
Treatment Group 3	Frequency	Percent	Cumulative Percent
Valid -3	1	5.6	5.6
-1	3	16.7	22.2
0	2	11.1	33.3
1	4	22.2	55.6
2	3	16.7	72.2
3	3	16.7	88.9
4	1	5.6	94.4
5	1	5.6	100.0
Total	18	100.0	
Control Group	Frequency	Percent	Cumulative Percent
Valid -3	1	3.0	3.0
-2	1	3.0	6.1
-1	5	15.2	21.2
0	11	33.3	54.5
1	6	18.2	72.7
2	3	9.1	81.8
3	3	9.1	90.9
4	1	3.0	93.9
5	1	3.0	97.0
6	1	3.0	100.0
Total	33	100.0	

Table 3. Frequency of pre- and post-test differences among 3rd graders

Treatment Group 1	Frequency	Percent	Cumulative Percent
Valid -3	1	5.3	5.3
-2	1	5.3	10.5
-1	2	10.5	21.1
0	2	10.5	31.6
1	1	5.3	36.8
2	1	5.3	42.1
3	3	15.8	57.9
4	3	15.8	73.7
5	2	10.5	84.2
6	2	10.5	94.7
9	1	5.3	100.0
Total	19	100.0	
Treatment Group 2	Frequency	Percent	Cumulative Percent
Valid -1	1	5.3	5.3
0	3	15.8	21.1
2	6	31.6	52.6
3	1	5.3	57.9
4	1	5.3	63.2
5	3	15.8	78.9
6	1	5.3	84.2
7	1	5.3	89.5
8	1	5.3	94.7
9	1	5.3	100.0
Total	19	100.0	
Control Group	Frequency	Percent	Cumulative Percent
Valid -5	1	2.6	2.6
-4	2	5.1	7.7
-3	1	2.6	10.3
-2	4	10.3	20.5
-1	3	7.7	28.2
0	5	12.8	41.0
1	3	7.7	48.7
2	8	20.5	69.2
3	5	12.8	82.1
4	6	15.4	97.4
5	1	2.6	100.0
Total	39	100.0	

Table 4. Frequency of pre- and post-test differences among 4th graders

Treatment Group 1	Frequency	Percent	Cumulative Percent
Valid -4	1	5.9	5.9
-3	2	11.8	17.6
-2	4	23.5	41.2
-1	1	5.9	47.1
1	2	11.8	58.8
3	3	17.6	76.5
5	2	11.8	88.2
10	1	5.9	94.1
11	1	5.9	100.0
Total	17	100.0	
Treatment Group 2	Frequency	Percent	Cumulative Percent
Valid -2	3	23.1	23.1
-1	1	7.7	30.8
0	1	7.7	38.5
1	2	15.4	53.8
2	1	7.7	61.5
4	2	15.4	76.9
6	1	7.7	84.6
8	2	15.4	100.0
Total	13	100.0	
Treatment Group 3	Frequency	Percent	Cumulative Percent
Valid 0	4	22.2	22.2
1	1	5.6	27.8
2	1	5.6	33.3
3	3	16.7	50.0
4	1	5.6	55.6
5	3	16.7	72.2
6	4	22.2	94.4
8	1	5.6	100.0
Total	18	100.0	
Treatment Group 4	Frequency	Percent	Cumulative Percent
-4	1	5.9	5.9
0	2	11.8	17.6
1	1	5.9	23.5
2	1	5.9	29.4
3	3	17.6	47.1
4	4	23.5	70.6
6	2	11.8	82.4
7	1	5.9	88.2
8	1	5.9	94.1
10	1	5.9	100.0
Total	17	100.0	

Control Group	Frequency	Percent	Cumulative Percent
Valid -5	2	5.1	5.1
-4	3	7.7	12.8
-3	4	10.3	23.1
-2	3	7.7	30.8
-1	5	12.8	43.6
0	2	5.1	48.7
1	2	5.1	53.8
2	3	7.7	61.5
3	3	7.7	69.2
4	3	7.7	76.9
5	1	2.6	79.5
6	2	5.1	84.6
7	2	5.1	89.7
9	3	7.7	97.4
11	1	2.6	100.0
Total	39	100.0	

Table 5. Frequency of pre- and post-test differences among 5th graders

Treatment Group 1	Frequency	Percent	Cumulative Percent
Valid 0	7	36.8	36.8
1	4	21.1	57.9
2	1	5.3	63.2
3	3	15.8	78.9
5	1	5.3	84.2
6	1	5.3	89.5
8	1	5.3	94.7
12	1	5.3	100.0
Total	19	100.0	
Treatment Group 2	Frequency	Percent	Cumulative Percent
Valid 0	1	8.3	8.3
2	1	8.3	16.7
3	4	33.3	50.0
4	2	16.7	66.7
5	2	16.7	83.3
7	1	8.3	91.7
9	1	8.3	100.0
Total	12	100.0	
Treatment Group 3	Frequency	Percent	Cumulative Percent
Valid 0	2	10.5	10.5
1	4	21.1	31.6
2	2	10.5	42.1
3	4	21.1	63.2
4	4	21.1	84.2
5	1	5.3	89.5

	6	1	5.3	94.7
	9	1	5.3	100.0
	Total	19	100.0	
Control Group		Frequency	Percent	Cumulative Percent
Valid	-4	2	6.3	6.3
	-3	1	3.1	9.4
	-2	2	6.3	15.6
	-1	7	21.9	37.5
	0	3	9.4	46.9
	1	4	12.5	59.4
	2	5	15.6	75.0
	4	6	18.8	93.8
	5	2	6.3	100.0
	Total	32	100.0	

Appendix E: Histograms

Table 1. Histogram compares frequency of gains among 1st grade participants

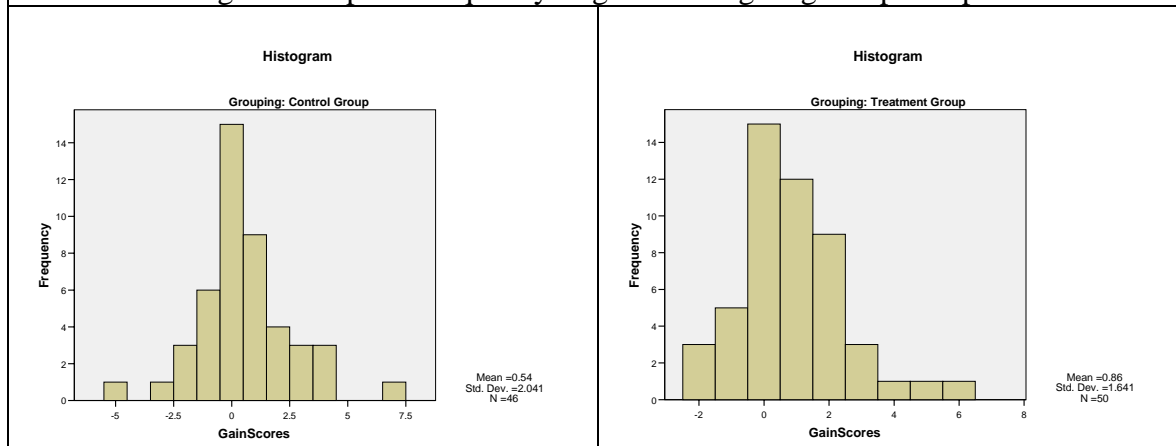


Table 2. Histogram compares frequency of gains among 2nd grade participants

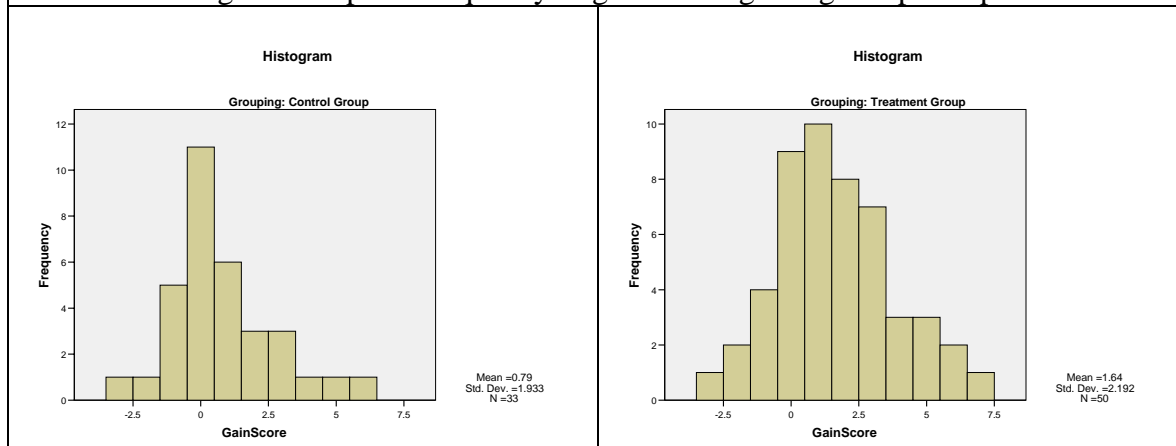


Table 3. Histogram compares frequency of gains among 3rd grade participants

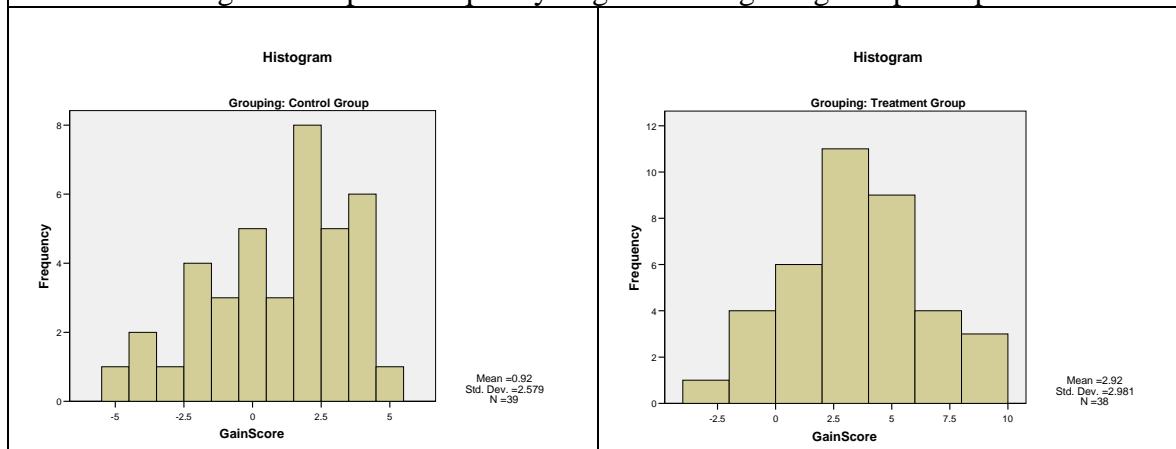


Table 4. Histogram compares frequency of gains among 4th grade participants

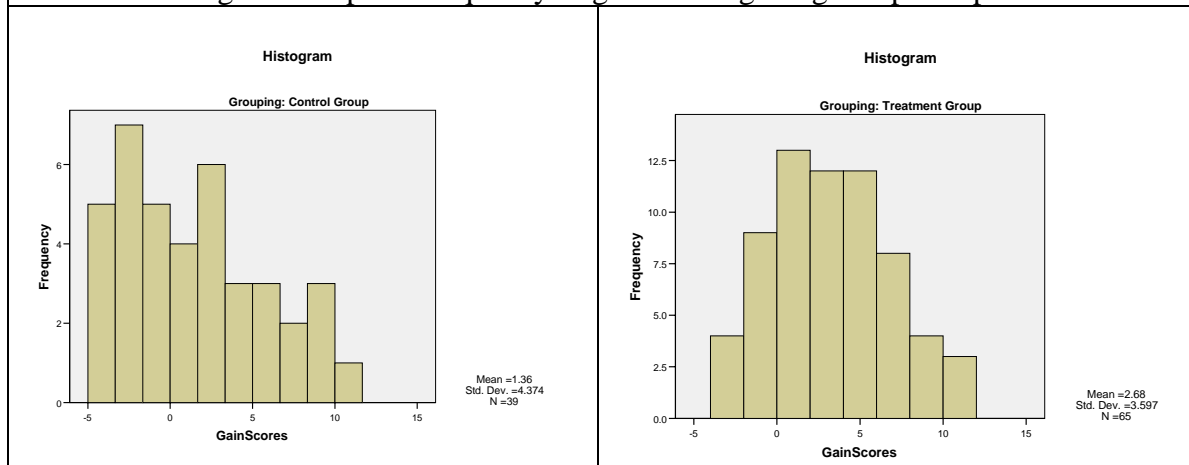
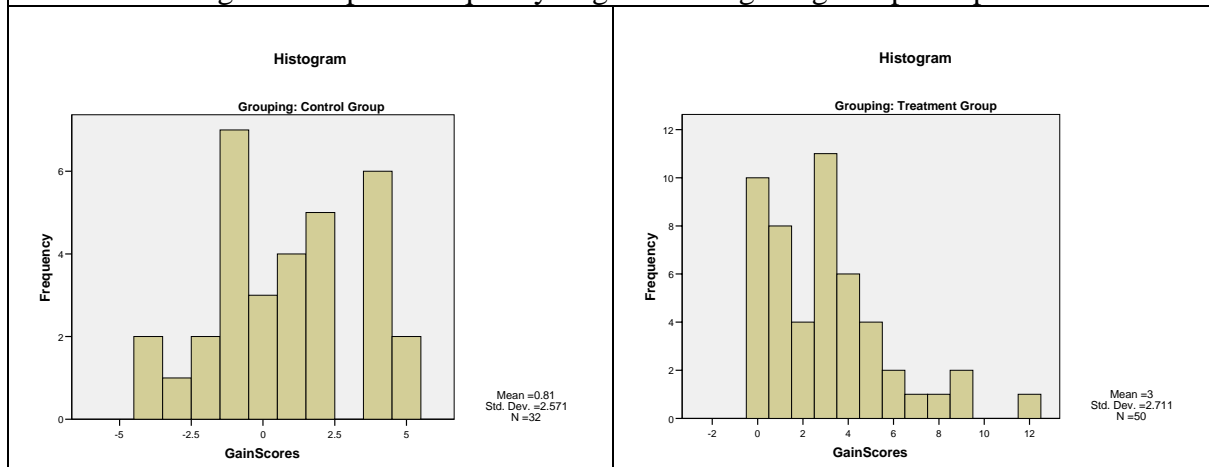


Table 5. Histogram compares frequency of gains among 5th grade participants



Appendix F: Exit Interview Questions

Teacher Name:

Date:

Post-Study Interview Questions: Treatment Teachers

1. Describe how you used the LLV curriculum with your students? (prompt for type of students, number of units completed, time spent)
2. What instructional strategies did you learn or strengthen while using the LLV curriculum?
3. What components of the LLV materials did you find the most useful? Least useful? Why?
4. What did you observe with regard to student engagement and motivation to learn with the LLV materials?
5. How did you use the LLV assessments?
6. Is the program easy to implement? What are some of the barriers to implementing it with small groups or individualized instruction?
7. Did you achieve what you hoped for with the LLV curriculum materials?
8. Any other questions/concerns that you are still wondering about with regard to the LLV curriculum?

Post-Study Interview Questions: Control Teachers

1. What are some of the instructional strategies that you generally use in your classroom?
2. How many ELL and at-risk students do you have? How many are reading at grade level?
3. What did you observe with regard to student engagement and motivation to learn with your curriculum materials?
4. What types of assessments do you use?
5. What do you like about your reading curriculum? Dislike about it?

Language, Literacy & Vocabulary Teacher Survey

There are three sections to this survey. Please respond to each question according to your experience during the spring pilot. This survey is being conducted by National Geographic to collect feedback from those involved in the LLV program pilot. Individual data are confidential. Names and other identifying data will be removed before reporting. No personal information will be shared with anyone. Data is used for research purposes only.

Section 1: Background Information

School: _____

Grade Level: _____

Name: _____

1. How many years have you been teaching reading in the content areas?

- a. ____ 1-2 years
- b. ____ 3-4 years
- c. ____ 5-6 years
- d. ____ 7+ years

2. Please identify the number of students in your classroom: _____

3. How many of the students in your classroom are currently:

- a. ____ new English language learners?
- b. ____ bridging English language learners?
- c. ____ at-risk readers in English?
- d. ____ at-risk readers in native language?
- e. ____ reading at grade level?

4. Which units did you use with students during the spring pilot? (Fill in the blank):

Section 2: Frequency of Use

5. During the spring pilot, on average, how often did you use the following LLV program components with students during a week? Place an "X" in the appropriate column for each component listed.

	Daily	3-4 Weekly	1-2 Weekly	Never
a. Audio Lesson				
b. Think and Discuss				
c. Theme Song				
d. Graphic Organizer				
e. Comprehension Master Book 1				
f. Activity Master Book 1				
g. Comprehension Master Book 2				
h. Activity Master Book 2				
i. Take Home Book				
j. Family Focus				

6. During the spring pilot, on average, how often did you use the following LLV assessments with students during a week? Place an "X" in the appropriate column for each component listed.

	Daily	3-4 Weekly	1-2X Weekly	Never
a. What I learned				
b. How I learned				
c. Oral Reading Record				
d. Retelling Guide				
e. Fluency Guide				
f. Writing Rubric				
g. Developmental Writing Checklist				
h. Content Vocabulary Checklist				
i. Oral Language Development Checklist				

7. What did you *like best* about the LLV program and why?

8. What did you *like least* about the LLV program and why?

Section 3: Instructional Approaches

9. The following statements represent instructional approaches for teaching reading in the content areas. Please <u>circle</u> the response that best indicates your level of engagement with each instructional approach during the LLV pilot.	<u>Level 0</u> I know little about this and do not plan to use it	<u>Level 1</u> I am learning about this but have not decided to use it	<u>Level 2</u> I am learning about this and plan to use it	<u>Level 3</u> I seldom practice this and am learning to do it better	<u>Level 4</u> I practice this sporadically with ease	<u>Level 5</u> I practice this regularly with confidence
Instructional Grouping Approaches:						
a. Whole group reading instruction	0	1	2	3	4	5
b. Small group reading instruction	0	1	2	3	4	5
c. Individualized reading instruction	0	1	2	3	4	5
d. Home-based reading activities	0	1	2	3	4	5
Reading Approaches:						
e. Developing oral language	0	1	2	3	4	5
f. Modeling comprehension strategies	0	1	2	3	4	5
g. Checking for understanding	0	1	2	3	4	5
h. Discussing the book	0	1	2	3	4	5
i. Rereading for fluency	0	1	2	3	4	5
j. Conducting think alouds	0	1	2	3	4	5
k. Developing key content concepts	0	1	2	3	4	5
l. Developing Key Vocabulary Words	0	1	2	3	4	5
Writing Approaches:						
m. Modeling writing	0	1	2	3	4	5
n. Scaffolding shared writing	0	1	2	3	4	5
o. Facilitating guided writing	0	1	2	3	4	5
Student Assessment Approaches:						
p. Conducting whole group assessment before unit	0	1	2	3	4	5
q. Conducting whole group assessment during unit	0	1	2	3	4	5
r. Conducting small group assessment during unit	0	1	2	3	4	5
s. Conducting individual assessment during unit	0	1	2	3	4	5
t. Conducting whole group assessment at end of unit	0	1	2	3	4	5

Language, Literacy & Vocabulary Teacher Survey

There are three sections to this survey. Please respond to each question according to your experience during the spring pilot. This survey is being conducted by National Geographic to collect feedback from those involved in the LLV program pilot. Individual data are confidential. Names and other identifying data will be removed before reporting. No personal information will be shared with anyone. Data is used for research purposes only.

Section 1: Background Information

School: _____

Grade Level: _____

Name: _____

1. How many years have you been teaching reading in the content areas?

- a. ____ 1-2 years
- b. ____ 3-4 years
- c. ____ 5-6 years
- d. ____ 7+ years

2. Please identify the number of students in your classroom: _____

3. How many of the students in your classroom are currently:

- a. ____ new English language learners?
- b. ____ bridging English language learners?
- c. ____ at-risk readers in English?
- d. ____ at-risk readers in native language?
- e. ____ reading at grade level?

4. Which units did you use with students during the spring pilot? (Fill in the blank):

Section 2: Frequency of Use

5. During the spring pilot, on average, how often did you use the following LLV program components with students during a week? Place an "X" in the appropriate column for each component listed.

	Daily	3-4 Weekly	1-2 Weekly	Never
a. Audio Lesson				
b. Build Background transparency				
c. Vocabulary transparency				
d. Comprehension Strategy Explanation transparency				
e. Comprehension Strategy Checklist transparency				
f. Graphic Organizers				
g. Study Guides				
h. Vocabulary Masters				
i. Comprehension Masters				
j. Language Masters				
k. Writing organizer				
l. Home-School Connection				

6. During the spring pilot, on average, how often did you use the following LLV assessments with students during a week? Place an "X" in the appropriate column for each component listed.

	Daily	3-4 Weekly	1-2X Weekly	Never
a. Learning Master Pre-Test				
b. Learning Master Post-Test				
c. Progress Tracking Form				
d. Student Self-Assessment				
e. Research and Write Rubric				

7. What did you *like best* about the LLV program and why?

8. What did you *like least* about the LLV program and why?

Section 3: Instructional Approaches

9. The following statements represent instructional approaches for teaching reading in the content areas. Please <u>circle</u> the response that best indicates your level of engagement with each instructional approach during the LLV pilot.	<u>Level 0</u> I know little about this and do not plan to use it	<u>Level 1</u> I am learning about this but have not decided to use it	<u>Level 2</u> I am learning about this and plan to use it	<u>Level 3</u> I seldom practice this and am learning to do it better	<u>Level 4</u> I practice this sporadically with ease	<u>Level 5</u> I practice this regularly with confidence
Instructional Grouping Approaches:						
a. Whole group reading instruction	0	1	2	3	4	5
b. Small group reading instruction	0	1	2	3	4	5
c. Individualized reading instruction	0	1	2	3	4	5
d. Home-based reading activities	0	1	2	3	4	5
Reading Approaches:						
e. Developing oral language	0	1	2	3	4	5
f. Modeling comprehension strategies	0	1	2	3	4	5
g. Checking for understanding	0	1	2	3	4	5
h. Discussing the book	0	1	2	3	4	5
i. Rereading for fluency	0	1	2	3	4	5
j. Conducting think alouds	0	1	2	3	4	5
k. Developing key content concepts	0	1	2	3	4	5
l. Developing Key Vocabulary Words	0	1	2	3	4	5
Writing Approaches:						
m. Modeling writing	0	1	2	3	4	5
n. Scaffolding shared writing	0	1	2	3	4	5
o. Facilitating guided writing	0	1	2	3	4	5
Student Assessment Approaches:						
p. Conducting whole group assessment before unit	0	1	2	3	4	5
q. Conducting whole group assessment during unit	0	1	2	3	4	5
r. Conducting small group assessment during unit	0	1	2	3	4	5
s. Conducting individual assessment during unit	0	1	2	3	4	5
t. Conducting whole group assessment at end of unit	0	1	2	3	4	5

Appendix I: Classroom Observation Protocol

Content Literacy Observation Protocol Directions

The Content Literacy Observation Protocol is designed for use in the National Geographic School Publishing validation studies. Its primary purpose is to document teaching and learning behaviors related to content literacy instruction. The Protocol is designed for OBSERVATION, NOT EVALUATION, of teaching and learning. Each protocol is intended for use in a single 15-20 minute observation.

A minimum of 1 and a maximum of 3 observations with each teacher (treatment and control) are needed for the validation study.

A principal, assistant principal, reading specialist, or reading coach can conduct the Content Literacy Observation.

Directions:

- Contact the teacher and request a time to come and observe. Be sure that you will observe a content literacy (science or social studies) lesson.
- Prior to the observation, complete the information at the top of the form (Teacher, Grade, School, Observer, and Date).
- For each OBSERVED item, place an “X” to indicate if the behavior was Observed (O) or Not Observed (N).
- NOT ALL sections of the Observation protocol must be completed during an observation. See the chart below for further information. If you DO NOT observe a component, DO NOT mark that section.

Component I: Classroom Setting	MUST be completed for each observation. Please observe the overall classroom setting, not just the group the teacher is working with at the time.
Component II: Before Reading	MAY be completed, depending on what portion of the lesson is observed.
Component III: During Reading	MAY be completed, depending on what portion of the lesson is observed.
Component IV: After Reading	MAY be completed, depending on what portion of the lesson is observed.
Component V: Strategy and Skill Instruction	SHOULD be completed for each observation, since strategy and skill instruction should be occurring before, during, and after reading.
Comments	MUST be completed for each observation to record overall impression of the observation.

- In the comments section, indicate if the lesson observed included before, during, and/or after reading instruction.
- Make 1 copy of the Observation Protocol and give it to the teacher.
- Send the original, by mail or fax, to National Geographic School Publishing at the address provided.