

Evaluation Study of Language, Literacy, & Vocabulary! Spring 2006 Pilot

submitted to:
Charlene Cobb
Director of Research and Curriculum
National Geographic School Publishing
990 Grove St., Suite 500
Evanston, IL 60201
847-864-1040
ccobb@ngs.org



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 Pilot Study	4
Methodology	4
Spring 2006 Pilot Results	5
LLV Classroom Adoption Practices	8
Recommendations for Fall 2006 Study	10
Appendices	
Appendix A: Test Validation	11
Appendix B: Descriptive statistics by grade level	13
Appendix C: Group frequency charts by grade level	15
Appendix D: Histograms	18
Appendix E: Focus Group Questions	22
Appendix F: WOL Survey	23
Appendix G: RE Survey	26

Executive Summary

Purpose

A pilot study of the *Language, Literacy, & Vocabulary!* program was conducted by Learning Gauge, Inc., for National Geographic School Publishing, in spring 2006. The program evaluation examined the classroom adoption approaches used by participating teachers and the subsequent impact of the *Language, Literacy, & Vocabulary!* (LLV) program on students' content area literacy development.

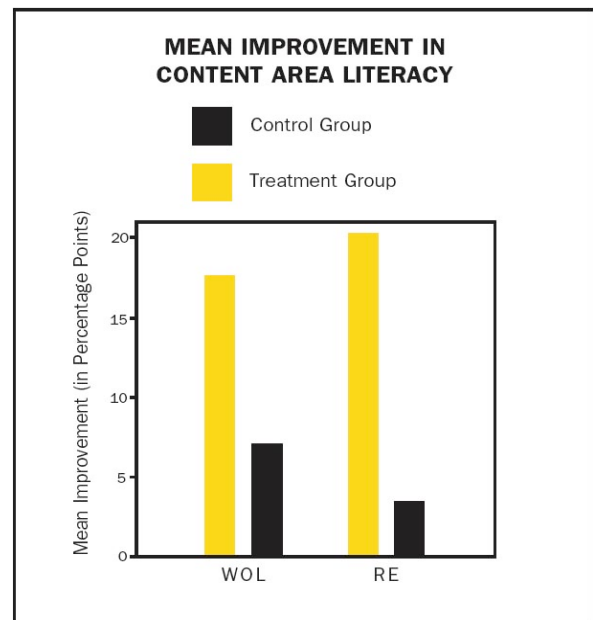
The LLV program provides curriculum resources for vocabulary and comprehension development through study of academic content. Materials are designed for multi-need instruction of English language learners, at-risk readers, and special need learners.

Methodology

The study was conducted in urban Illinois public schools with 259 students in 4 schools, 90 students in the treatment groups and 169 students in the control groups. Classes of students were assigned to treatment or control groups. Pre- and post-test data from students provides insight into how effectively the *Language, Literacy, & Vocabulary!* program improves students' content area literacy. Teachers in the treatment classrooms taught 3 to 6 units and used a wide variety of LLV instructional tools and assessments with students on a daily basis. Weekly units were taught in a continuous duration.

Results

A thorough analysis of the data collected indicates that the students in treatment groups made statistically significant gains in their content area literacy development compared with student who did not use the *Language, Literacy, & Vocabulary!* curriculum. In each curriculum group, the mean improvement from pretest to posttest was substantially higher for the students in the treatment groups. In addition to gains on the pre- and post-test, student groups who used the *Language, Literacy, & Vocabulary!* program showed considerably less variance and were more consistent in their literacy growth compared with the wide ranging variance, including frequent negative gains, among control group students.



Purpose of the LLV Pilot Study

A pilot study of the *Language, Literacy, & Vocabulary!* program was conducted by Learning Gauge, Inc., for National Geographic School Publishing, in spring 2006. The program evaluation examined the classroom adoption approaches used by participating teachers and the subsequent impact of the *Language, Literacy, & Vocabulary!* (LLV) program on students' content area literacy development.

The LLV program provides curriculum resources for vocabulary and comprehension development through study of academic content. Materials are designed for multi-need instruction of English language learners, at-risk readers, and special needs learners.

Methodology

The evaluation for the LLV involved a two-fold design. First, a validation study of the criterion-referenced tests developed to measure students' academic vocabulary and reading comprehension strategies was conducted (see Appendix A). Second, the pre- and post-test results of the validated items were analyzed. Additional data were gathered from treatment group teachers about their classroom adoption practices with regard to the *Language, Literacy & Vocabulary!* program. These data allowed for analysis of teachers' program implementation logs, a post-implementation survey, pre- and post-implementation focus groups, and samples of student work. Pre- and post-test data from students provided insight into how effectively the LLV implementation improved students' content area literacy.

The evaluation was conducted in urban Illinois public schools with 292 students in 4 schools, 123 students in the treatment groups and 169 students in the control groups. Classes of students were assigned to treatment or control groups. Teachers from the treatment classrooms received a half-day professional development orientation to the program, participated in pre- and post-treatment focus groups, and completed a survey and implementation logs about their curriculum usage. The number of LLV units treatment group teachers implemented ranged from 1 to 6 units while teachers in the control classrooms taught their regular reading and content area curricula. At the beginning and conclusion of the pilot, students were administered a grade level criterion-referenced test. Students also completed formal and informal assessments that were part of the LLV program, although assessment usage varied among teacher groups.

Characteristics of Treatment Group Participants

The majority of the treatment group students were bridging English language learners at the end of the school year. The treatment groups also included new English language learners, at-risk readers in English, at-risk readers in native language and a small percent of grade level readers.

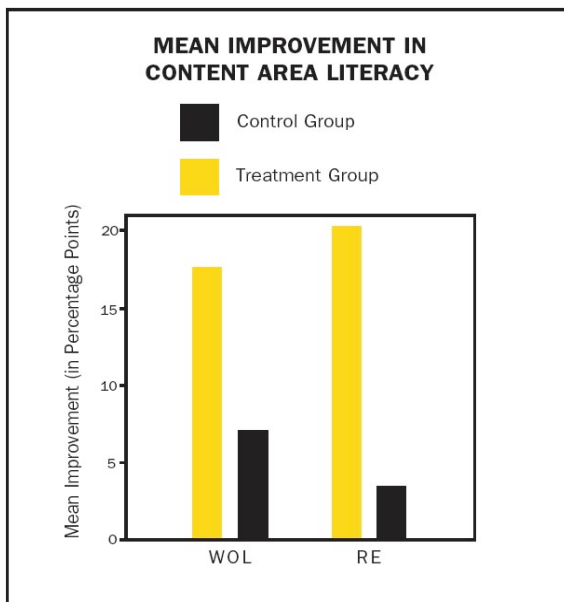
The One-sample Kolmogorov-Smirnov (1-Sample K-S) test was conducted on all of the grade level groupings. The 1-Sample K-S results showed all participant groups fell

within the normal population distribution except for Grade 1 which had some outliers. Once the outliers were dismissed, the test results did show it was a normal population. The Levene Test of Homogeneity of Variances also was conducted on each of the grade level pre-test group data to determine if the data met the requirements for comparability of the groups. The results of the Levene Test indicated that all of the grade level groups are indeed comparable. The sample population, although normal, did show a wide range and variability on pre-test performance. This is not surprising since the learners involved in the pilot study represent a diverse population of literacy backgrounds in native foreign languages and English as a second language.

Results of the Pilot Study

A Paired Samples T-Test was conducted on the validated pre- and post-tests for Grades 1, 2, 3, 4, and 5. For purposes of Table 1, only those teachers who taught 3 to 6 units and used a wide variety of LLV instructional tools and assessments for four to five days per week were included in the summary of results. These teachers also indicated teaching the weekly units in a continuous duration. Other teachers in treatment classrooms indicated their implementation was weak with less than 3 units implemented sporadically rather than on a daily basis or continuous weeks. These latter treatment classes were excluded from the results represented in Table 1 as their access to the LLV curriculum was limited.

Table 1. Summary of spring 2006 pilot results.



In Table 2, the statistical results include data from all eight treatment classes. Significant increases in students' content area literacy development were found for 5 of 8 treatment groups. The analysis of means in Table 2 shows a strong teacher effect. Additional statistical tables comparing means, frequency of pre- and post-test differences and histograms are compiled in the Appendices B-D.

Table 2. Results of the Paired Samples T-Test measuring differences between pre- and post-test performances for participants in the LLV pilot.

	1st Grade Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
				Lower	Upper			
1st Grade Treatment Group 1 n = 17	1.059	1.298	.315	.392	1.726	3.364*	16	.004
1st Grade Treatment Group 2 n = 11	2.273	2.240	.675	.768	3.778	3.365*	10	.007
1st Grade Control Group n = 36	.583	1.339	.223	.130	1.036	2.614*	35	.013
2nd Grade Paired Differences								
2nd Grade Treatment Group 1 n = 19	.105	1.761	.404	-.743	.954	.261	18	.797
2nd Grade Treatment Group 2 n = 20	1.650	1.309	.293	1.037	2.263	5.638*	19	.0001
2nd Grade Treatment Group n = 42	.357	1.590	.245	-.138	.853	1.456	41	.153
3rd Grade Paired Differences								
3rd Grade Treatment Group n = 12	1.500	1.679	.485	.433	2.567	3.095*	11	.010
3rd Grade Control Group n = 38	.132	1.711	.278	-.431	.694	.474	37	.638
4th Grade Paired Differences								
4th Grade Treatment Group n = 8	-.250	.886	.313	-.991	.491	-.798	7	.451
4th Grade Control Group n = 33	.182	1.740	.303	-.435	.799	.600	32	.553
5th Grade Paired Differences								
5th Grade Treatment Group 1 n = 11	1.091	1.640	.495	-.011	2.193	2.206*	10	.052
5th Grade Treatment Group 2 n = 25	-.120	1.364	.273	-.683	.443	-.440	24	.664
5th Grade Control Group n = 20	.250	1.164	.260	-.295	.795	.960	19	.349

* indicates significant difference between pre- and post-test performance

Both of the teacher class groups in the 1st grade treatment show a significant gain. The Paired Samples T-Test results for the 1st Grade treatment groups are $t(16) = 3.364$ and $t(10) = 3.365$; $p < 0.05$ which indicates that the 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 $t(35) = 2.614$; $p < 0.05$, which also indicated a significant increase in pre- and post-test performance. The control group increase may have been influenced by other reading interventions and/or the read aloud procedure used to administer the test. Nonetheless, both treatment groups' pre-test means started below and then surpass that of the control groups. The second treatment group showed the most overall gain from a pre-test mean of 6.18 to post-test mean of 8.45. For the first treatment group the pre-test mean was 7.41 and post-test mean was 8.47. The second treatment group also had less variance than the first treatment group although the class size differed substantially. Overall, 68% of students in the treatment groups showed some gain while 50% of

students in the control group did not show any gain. In addition, 18% of students in the treatment groups gained 1 or 2 points more than students in the control group. Frequency charts and histograms comparing differences in students' pre- and post-test performance are compiled in Appendix C and D, respectively.

The Paired Samples T-Test result for the 2nd Grade treatment group 1 is $t(18) = .261$; $p < 0.05$ which indicates no significance between pre- and post-test performance in this class group. The teacher survey for this class indicates the LLV program was used weekly however some of the student assessments were not used. The Paired Samples T-Test result for the 2nd Grade treatment group 2 is $t(19) = 5.638$; $p < 0.05$ which indicates a significant learning gain among students in this class group. This teacher did not submit implementation logs but reported during the post-implementation focus group that he had made extensive use of the LLV program set up in learning centers for students. The control group's Paired Samples T-Test result is $t(41) = 1.456$; $p < 0.05$, which does not indicate a significant difference in students' pre- and post-test performance. The mean for students in the 2nd grade treatment group 2 started below the mean of the control group and then surpassed it on the post-test. The comparison of means and frequency of pre- and post-test differences indicates a clear teacher effect for the treatment group classes. Treatment group 1 performed similar to the control group. In contrast, treatment group 2 had substantial gains with 85% of students increasing their score in a range from 1-4.

The Paired Samples T-Test result for the one 3rd Grade treatment group is $t(11) = 3.095$; $p < 0.05$ 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 $t(37) = .474$; $p < 0.05$, which indicates no significant difference in pre- and post-test performance. In addition, the treatment group's mean started below and then surpassed that of the control group. Seventy-five percent of students in the treatment group showed some gain while only 34% of students in the control group showed any gain during the period of instruction. In addition, 33% of students in the treatment group gained 3 or more points while only 8% of students in the control group showed a gain of 3 or more points.

The Paired Samples T-Test result for the one 4th Grade treatment group is $t(7) = -.798$; $p < 0.05$ which indicates that students' post-test scores were considerably below their pre-test performance. These data are somewhat unreliable, however, as the power of the test was weak (see Appendix A) and their teacher's implementation of the LLV inconsistent.

The Paired Samples T-Test for 5th Grade treatment group 1 is $t(10) = 2.206$; $p < 0.05$ which indicates that these students performed significantly better on the post-test than they did on the pre-test. The results for treatment group 2 is $t(24) = -.440$; $p < 0.05$ and for control group $t(19) = .960$; $p < 0.05$. Both of the latter results do not indicate a significant different in students' pre- and post-test performance. Students in treatment group 2 performed on a par with students in the control group, both of which had 35% gains on their pre- and post-test performance. Meanwhile, 73% of students in treatment group 1 showed gains ranging from 1 to 3 with 27% of those gaining 3 points. Here again we see a strong teacher effect, which is discussed in more detail in the following section of this report.

Overall, students in treatment groups showed a dramatic decrease in variance (see Appendix B). This phenomenon indicates treatment students whose teachers fully implemented the program became more consistent in their content area literacy performance during the LLV pilot. Meanwhile, the control groups' performance is inconsistent with many students showing negative gains (see frequency tables in Appendix C). Another finding from the pilot study is that of teacher effects. Thus class groups within the treatment groups will be an important factor to track closely during the fall 2006 LLV study. It is also interesting to note that the treatment classes showing significant gains had a pre-test mean below that of their control group and then surpassed the post-test performance of students in control classrooms.

LLV Classroom Adoption Practices

Teachers in the treatment groups took a variety of classroom adoption approaches. Table 3 shows the details from implementation data collection sources such as implementation logs, teacher post-surveys, focus groups and submitted student work samples.

Analysis of these data indicates that teachers who used a wide variety of program materials and who implemented *at least* 3 units daily for a continuous 3-week duration had students who performed significantly better than the students in control groups. The exception to this is the 2nd grade teacher from Group 1. The reason for a non-significant result with this group is unclear. She was one of three teachers who had only been teaching reading in the content areas of 1-2 years. However, the other six participating teachers had 5 or more years of experience teaching reading in the content areas and the 4 and 5 grade teachers whose students also had no significant gains had 7 or more years of teaching experience. Therefore, it is not possible to make conclusions based on prior teaching experience. The teachers in the successful treatment groups also implemented assessments that provided students with formative feedback about their learning progress. This was evident in the groups where teachers used the LLV units' pre- and post-tests or informal assessments that were graded for individual students. Teachers in treatment classrooms that did not show significant growth in content area literacy may have used a post-test or an informal assessment; however, there was no evidence of providing feedback to students during the instructional timeframe.

Table 3. Shows the range of implementation among teachers in treatment groups.

Treatment Group	student gains	units of study	times per week/Duration	assessment feedback evidence	LLV components*
1st grade					
Group 1	significant	5	5x/5wk	yes	AL; LM; ASMT;CL
Group 2	significant	6	5x/6wk	yes	AL; LM; ASMT;CL
2nd grade					
Group 1	non-significant	3	5x/3wk	yes	AL; LM; ASMT;CL
Group 2	significant	3	no logs available	yes	AL; LM; ASMT in learning centers
3rd grade					
Group 1	significant	6	5-6x/6wk	yes	AL; TR; LM; ASMT pre- & post-tests
Group 2	reason for failure to implement unknown				
4th grade					
Group 1	non-significant	4	3-4x/6wk	no	TR & LM inconsistency across units; ASMT post-tests only;
Group 2	teacher unable to implement LLV program due to other curriculum priorities				
5th grade					
Group 1	significant gains	4	4-5x/4wk	yes	TR; LM; ASMT pre- & post-tests
Group 2	non-significant	2	2x/5wk	no	TR; LM; ASMT post-tests only;

* LLV Codes [WOL AL = Audio Lesson; LM = Learning Masters; ASMT = Assessments; CL Checklists] [RE AL = Audio Lesson; TR = Transparencies; LM = Learning Masters; ASMT = Assessments]

Nine teachers attended the post-implementation focus group sessions and completed the post-survey. All nine teachers rated themselves regular and confident practitioners of whole group and small group reading instruction. The 1st and 2nd grade teachers also said they regularly practice individualized and home-based reading activities with confidence. Most of the 1st and 2nd grade teachers also indicated regular and confident practice with a variety of student assessment approaches including pre- and post-testing and informal assessment with whole class group, small groups, and individual students. The majority of 3rd, 4th and 5th grade teachers said they either seldom or sporadically practice individualized and/or home-based reading activities. In addition, many of the 3rd, 4th and 5th grade teachers indicated either a lack of or only sporadic practice with a variety of student assessment approaches including pre-testing and informal assessment with whole class group, small groups, and individual students. The majority of all of the teachers in the treatment groups indicated a lack of regular practice with writing approaches. Only 66% of the teachers rated their practice as regular and confident with regard to helping students develop oral language and key content area concepts. Only 55% of teachers regularly practice developing students' key vocabulary words.

Teachers overwhelmingly agreed, in the focus groups, that the LLV materials foster student engagement and interest more than other curriculum resources. Since learner

engagement and interest contribute to students' developing ownership of literacy, which leads to self-motivated literacy development habits, the content and graphical appeal of these materials is a contributing factor in the success found among treatment groups those teachers fully implemented the program.

Recommendations for the Fall Study

The following are recommendations for the fall 2006 study, based on the findings from the spring 2006 pilot:

- Validity of the tests can be strengthened for the fall by strengthening the power of the test with a larger sample of students and adding 3-4 test items for each unit's vocabulary and 3-4 items for each comprehension strategy to be measured. Since many factors were found in the tests used in the pilot, it is also recommended that the tests for the fall study represent a stronger alignment between LLV units implemented in the treatment groups and test item content, particularly for the vocabulary sections.
- Information about the types of learners involved in this study was limited so it would be helpful to add a student code to the pre- and post-tests. Teachers in both treatment and control groups would be asked to complete the student coding with the test administrator. Implementation data from teachers also need to be strengthened. Suggestions for adjustments to the implementation data collection procedures will be shared verbally with the director of research at NGS.
- The strong teacher effect in the pilot data indicates a need for more professional development to strengthen implementation of the LLV program, based on effectiveness parameters found. That is, teachers who used the full variety of LLV program components and taught 3 or more units on a daily basis tended to have students who showed significant growth in their literacy development. Teachers also varied in their understanding of program approaches particularly in the areas of assessment, how to develop students' key content concepts, and home-based activities.

Appendix A: Test Validation

A validation study was conducted using the pre-test data from the participating schools involved in the spring 2006 pilot for the Language, Literacy and Vocabulary program. Learning Gauge conducted statistical analyses using classical methods for establishing the validity and reliability of criterion-referenced tests. The validation measures included the following statistical analysis steps for each grade level test:

- Determining the normality of the sample population using the One-Sample Kolmogorov-Smirnov Test
- 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 whole test as well as 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 correlations between the vocabulary and comprehension subscales.
- Conducting the Cronbach Alpha test for reliability.
- Generating basic statistics such as analysis of means, standard deviation, effect size, and standard error of measurement.
- Regression analysis.

This initial validation study of the criterion-referenced tests indicated low validity and reliability for some of the test items. Therefore, a second tier validation study was conducted to generate a subscale of test items that represent the strongest reliability and validity for the pre- and post-test analysis of student learning results. The second tier validation process for these item scales is explained in Tables 1-3 below. The pre- and post-test results in the report are based on the most dependable test scales available for each grade level (see Table 1 for specific items).

Table 1 shows the Cronbach Alpha for each of the revised measurement scales. These tests have mid-level reliability and could be improved for use in fall 2006. Also listed in Table 2 are the test items that make up the measurement scale used for the pre- and post-test analysis. Table 2 shows the Pearson Correlation for each of these revised grade level tests. In addition, Table 3 shows the power for each test used in the analysis.

For the fall study, it is recommended that 3-4 vocabulary items per unit and 3-4 items per comprehension strategy be used in each grade level test to strengthen the validity and reliability of the measurement scales. The planned increase in sample size will also provide stronger results upon which to make claims about the effectiveness of the *Language, Literacy, & Vocabulary!* program.

Table 1. Test Reliability Statistics for each of the grade level tests used for the pre- and post-test data analysis.

Grade Level	Cronbach's Alpha*	Test Items
First Grade	.557	Nine items (Questions 1, 2, 6-12)
Second Grade	.607	Five items (Questions 2, 5, 6, 7, 9)
Third Grade	.606	Eight items (Questions 1-4, 6, 9, 11, 15)
Fourth Grade	.510	Five items (Question 2, 4, 5, 12, 13)
Fifth Grade	.485	Five items (Question 1, 4, 8, 10, 13)

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

Table 2. Paired Samples Pearson Correlations

		N	Correlation	Sig.
First Grade	Treatment Group 1 & 2	28	.266	.172
	Control Group	36	.461	.005
Second Grade	Treatment Group 1	19	.268	.267
	Treatment Group 2	20	.533	.015
	Control Group	42	.157	.322
Third Grade	Treatment Group 1	12	.651	.022
	Control Group	38	.550	.0001
Fourth Grade	Treatment Group 1	8	.532	.175
	Control Group	33	.246	.168
Fifth Grade	Treatment Group 1	11	-.199	.558
	Treatment Group 2	25	.415	.039
	Control Group	20	.507	.023

Table 3. Power of the test for the Paired Samples T-Test (1-tailed).

Grade Level	Power of Test
1st Grade	0.755
2nd Grade	0.945
3rd Grade	0.773
4th Grade	0.247
5th Grade	0.549

Appendix B: 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	17	4	9	7.41	1.543	2.382
Post-Test	17	6	9	8.47	1.068	1.140
Treatment Group 2						
Pre-Test	11	3	9	6.18	1.618	2.618
Post-Test	11	7	9	8.45	.688	.473
Control Group						
Pre-Test	36	4	9	7.44	1.443	2.083
Post-Test	36	6	9	8.03	1.055	1.113
2nd Grade Descriptive Statistics						
	N	Minimum	Maximum	Mean	Std. Deviation	Variance
Treatment Group 1						
Pre-Test	19	0	5	3.74	1.485	2.205
Post-Test	19	1	5	3.84	1.425	2.029
Treatment Group 2						
Pre-Test	20	0	5	2.60	1.501	2.253
Post-Test	20	1	5	4.25	1.118	1.250
Control Group						
Pre-Test	42	1	5	3.52	1.234	1.524
Post-Test	42	2	5	3.88	1.214	1.473
3rd Grade Descriptive Statistics						
	N	Minimum	Maximum	Mean	Std. Deviation	Variance
Treatment Group 1						
Pre-Test	12	0	6	3.42	2.109	4.447
Post-Test	12	2	8	4.92	1.881	3.538
Control Group						
Pre-Test	38	0	7	3.47	1.899	3.607
Post-Test	38	0	6	3.61	1.685	2.840
4th Grade Descriptive Statistics						
	N	Minimum	Maximum	Mean	Std. Deviation	Variance
Treatment Group 1						
Pre-Test	8	1	4	2.63	.916	.839
Post-Test	8	1	4	2.38	.916	.839
Control Group						
Pre-Test	33	0	5	1.70	1.425	2.030
Post-Test	33	0	5	1.88	1.409	1.985

Fifth Grade Descriptive Statistics						
	N	Minimum	Maximum	Mean	Std. Deviation	Variance
Treatment Group 1						
Pre-Test	11	0	4	1.91	1.300	1.691
Post-Test	11	2	4	3.00	.775	.600
Treatment Group 2						
Pre-Test	25	0	5	3.12	1.394	1.943
Post-Test	25	1	5	3.00	1.080	1.167
Control Group						
Pre-Test	20	0	5	2.20	1.152	1.326
Post-Test	20	1	5	2.45	1.191	1.418

Appendix C: 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	0	6	35.3	35.3
	1	8	47.1	82.4
	2	1	5.9	88.2
	3	1	5.9	94.1
	5	1	5.9	100.0
	Total	17	100.0	
Treatment Group 2		Frequency	Percent	Cumulative Percent
Valid	-1	1	9.1	9.1
	0	2	18.2	27.3
	1	2	18.2	45.5
	2	1	9.1	54.5
	3	1	9.1	63.6
	4	1	9.1	72.7
	5	3	27.3	100.0
	Total	11	100.0	
Control Group		Frequency	Percent	Cumulative Percent
Valid	-2	3	8.3	8.3
	-1	3	8.3	16.7
	0	12	33.3	50.0
	1	9	25.0	75.0
	2	6	16.7	91.7
	3	3	8.3	100.0
	Total	36	100.0	

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

Treatment Group 1		Frequency	Percent	Cumulative Percent
Valid	-3	3	15.8	15.8
	-2	1	5.3	21.1
	0	7	36.8	57.9
	1	4	21.1	78.9
	2	3	15.8	94.7
	3	1	5.3	100.0
	Total	19	100.0	
Treatment Group 2		Frequency	Percent	Cumulative Percent
Valid	-1	1	5.0	5.0
	0	2	10.0	15.0
	1	7	35.0	50.0
	2	5	25.0	75.0
	3	3	15.0	90.0
	4	2	10.0	100.0
	Total	20	100.0	
Control Group		Frequency	Percent	Cumulative Percent
Valid	-3	3	7.1	7.1
	-2	2	4.8	11.9
	-1	7	16.7	28.6
	0	9	21.4	50.0
	1	9	21.4	71.4
	2	10	23.8	95.2
	3	2	4.8	100.0
	Total	42	100.0	

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

Treatment Group		Frequency	Percent	Cumulative Percent
Valid	-2	1	8.3	8.3
	0	2	16.7	25.0
	1	3	25.0	50.0
	2	2	16.7	66.7
	3	3	25.0	91.7
	4	1	8.3	100.0
	Total	12	100.0	
Control Group		Frequency	Percent	Cumulative Percent
Valid	-4	1	2.6	2.6
	-2	6	15.8	18.4
	-1	4	10.5	28.9
	0	14	36.8	65.8
	1	6	15.8	81.6
	2	4	10.5	92.1
	3	2	5.3	97.4
	5	1	2.6	100.0
	Total	38	100.0	

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

Treatment Group		Frequency	Percent	Cumulative Percent
Valid	-1	4	50.0	50.0
	0	2	25.0	75.0
	1	2	25.0	100.0
	Total	8	100.0	
Control Group		Frequency	Percent	Cumulative Percent
Valid	-4	1	3.0	3.0
	-3	2	6.1	9.1
	-1	9	27.3	36.4
	0	7	21.2	57.6
	1	6	18.2	75.8
	2	6	18.2	93.9
	3	1	3.0	97.0
	4	1	3.0	100.0
	Total	33	100.0	

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

Treatment Group 1- OC		Frequency	Percent	Cumulative Percent
Valid	-2	1	9.1	9.1
	-1	1	9.1	18.2
	0	1	9.1	27.3
	1	4	36.4	63.6
	2	1	9.1	72.7
	3	3	27.3	100.0
Total		11	100.0	
Treatment Group 2 - CH		Frequency	Percent	Cumulative Percent
Valid	-3	1	4.0	4.0
	-2	2	8.0	12.0
	-1	8	32.0	44.0
	0	5	20.0	64.0
	1	7	28.0	92.0
	2	1	4.0	96.0
	3	1	4.0	100.0
	Total		25	100.0
Control Group		Frequency	Percent	Cumulative Percent
Valid	-2	1	5.0	5.0
	-1	4	20.0	25.0
	0	8	40.0	65.0
	1	3	15.0	80.0
	2	4	20.0	100.0
	Total		20	100.0

Appendix D: Histograms

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

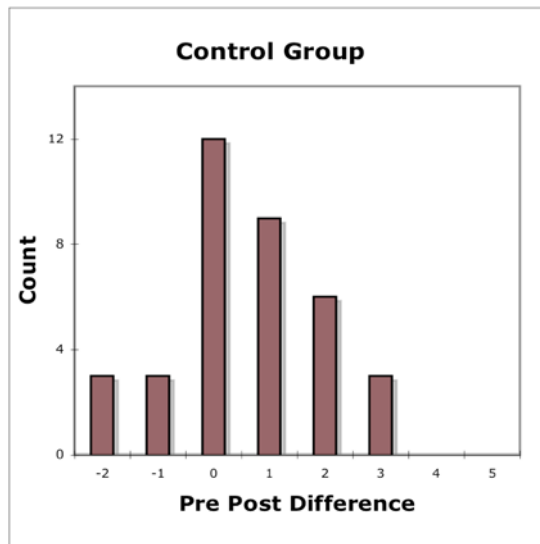
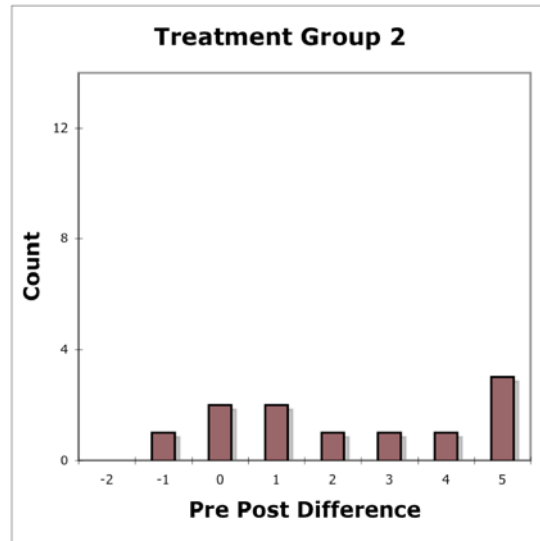
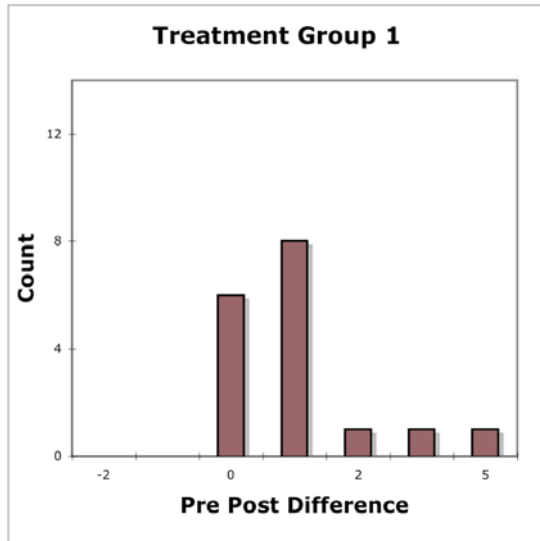


Table 2. Histogram compares frequency of gains among 2nd grade participants.

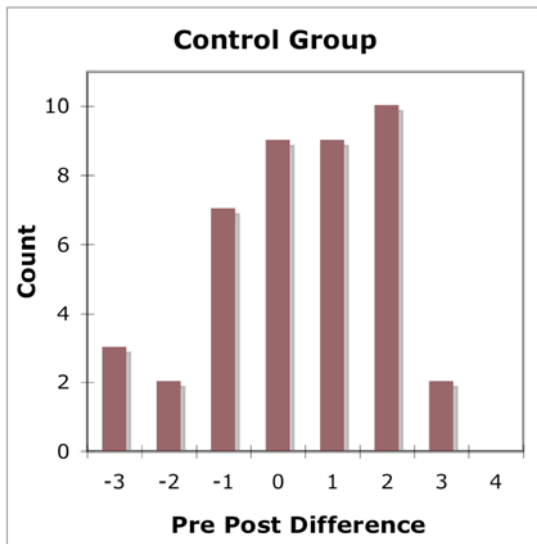
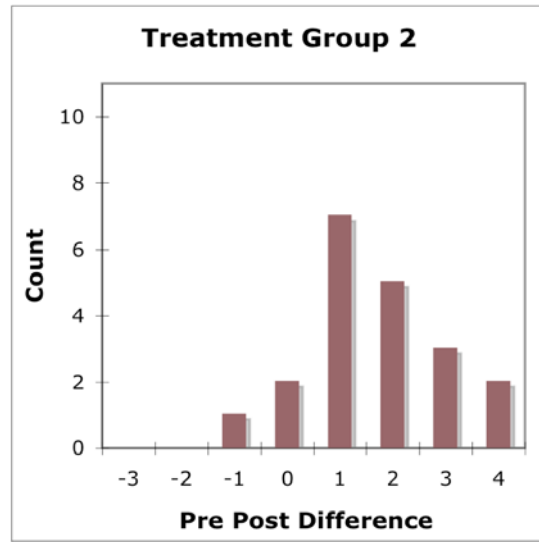
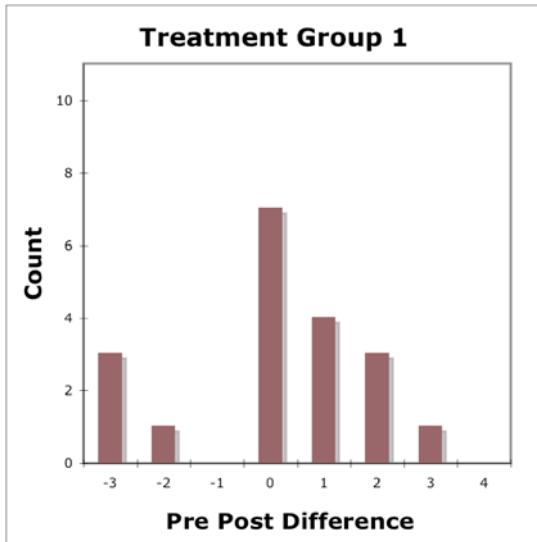


Table 3. Histogram compares frequency of gains among 3rd grade participants.

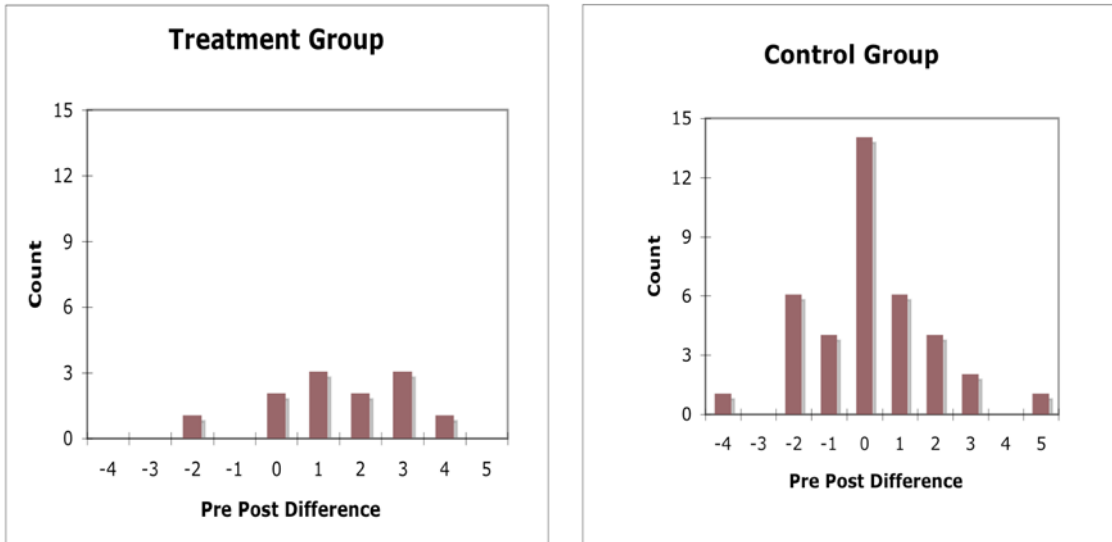


Table 4. Histogram compares frequency of gains among 4th grade participants.

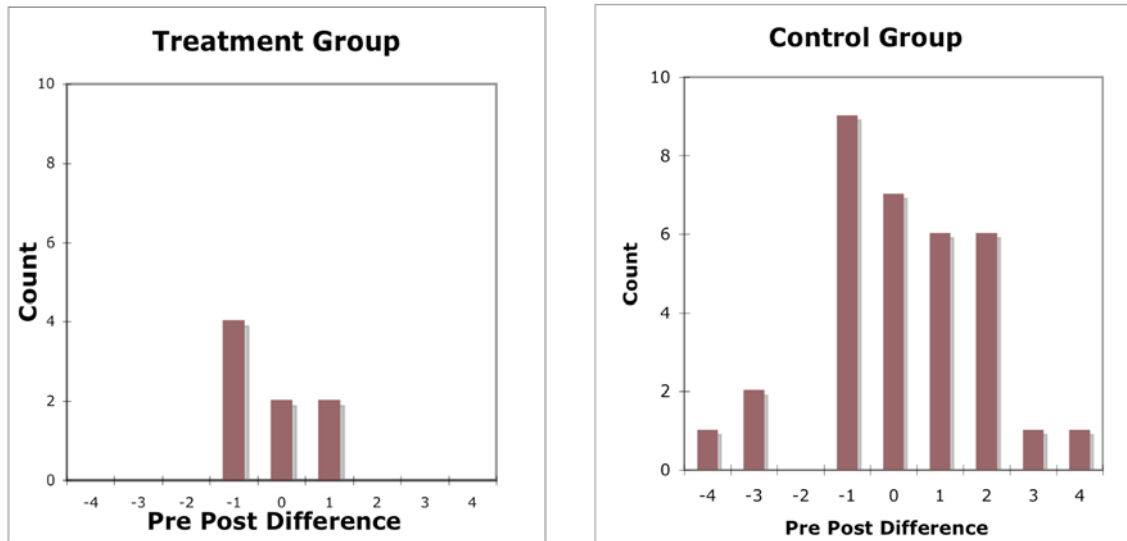
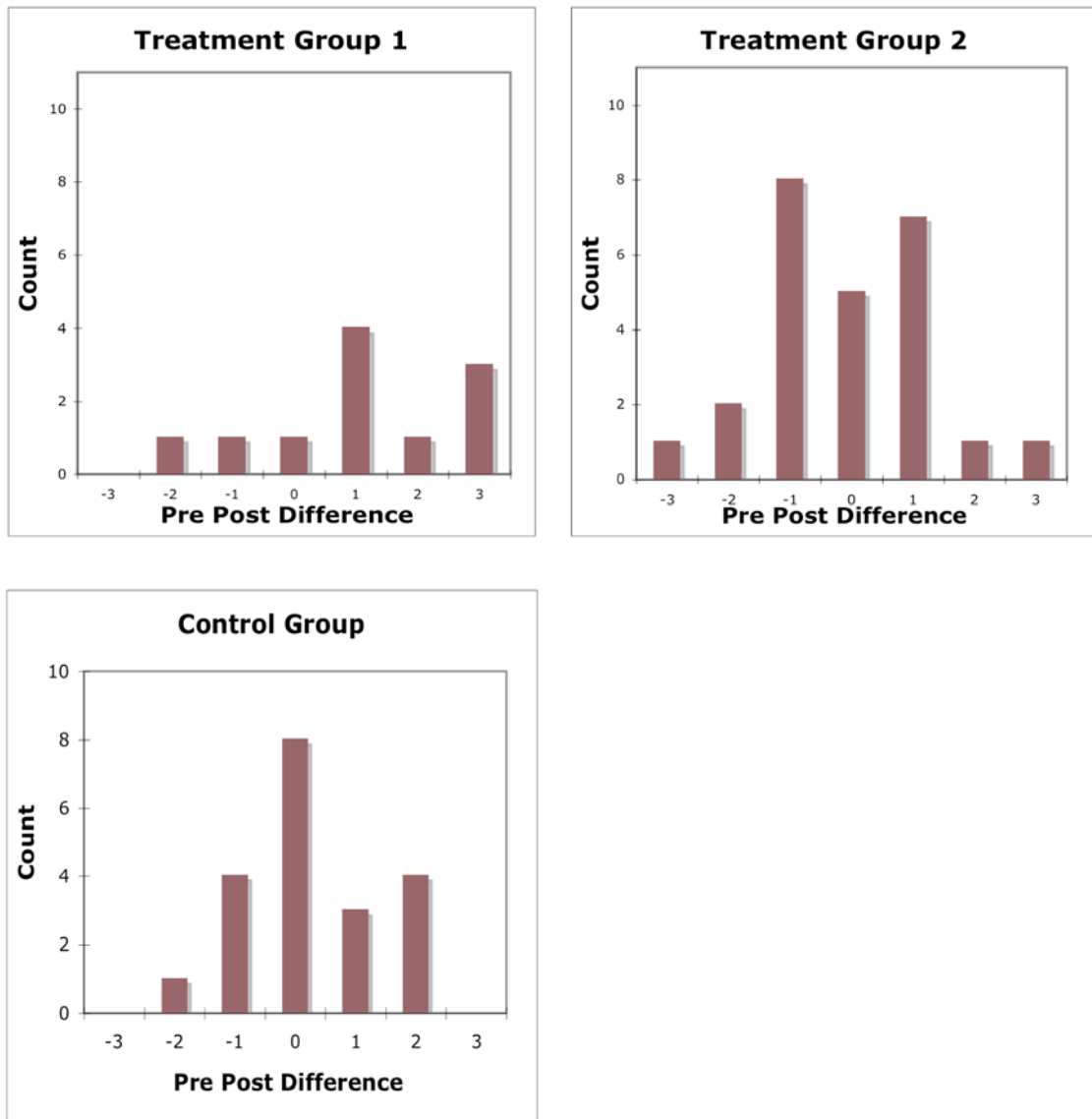


Table 5. Histogram compares frequency of gains among 5th grade participants.



Appendix E: Focus Group Questions

1. What are some of the instructional strategies for reading in the content areas that you learned or strengthened while using the LLV?
2. What are some of the instructional strategies for vocabulary building in the content areas that you learned or strengthened while using the LLV?
3. What components of the LLV materials did you find the most useful? Why?
4. How does the LLV materials compare with your other curriculum programs?
5. What did you observe with regard to student engagement and motivation to learn with these materials?
6. Do you feel that when your students exit the bi-lingual classroom, they will be more adequately prepared for science and social studies in the mainstream classroom because they have used these materials? Why or why not?
7. Is the program easy to implement? What are some of the barriers to implementing it with small groups or individualized instruction?
8. Did you achieve what you hoped for with the LLV materials?
9. Any other questions/concerns that you are still wondering about?

Appendix F: WOL Survey

Language, Literacy & Vocabulary Pilot 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 circle 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 G: RE Survey

Language, Literacy & Vocabulary Pilot 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 circle 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