An Evaluation of the *SRA FLEX Literacy* Program: A Pre-Publication Version Shows Promise

Shannon Flaum-Horvath, SKF Educational Services
Nancy E. Marchand-Martella, Ronald C. Martella, and Charalambos C. Cleanthous, Eastern Washington University

Academic instructional programs that are designed to promote higher-level literacy skills should be field-tested during the product development stage to determine their effectiveness. The purpose of this evaluation was to examine the effects of a pre-publication version of a digital-, print-, and project-based literacy program. A second purpose was to determine teacher satisfaction with the program. Results showed a mean Lexile® growth of 166.30L across all sites; 48% of the students exceeded the expected Lexile® growth from fall to spring assessments. On average, 29% of the expected yearly lessons were completed; the correlation between Lexile® growth and percentage of yearly lessons completed for the program was statistically significant. Finally, teachers reported satisfaction with the program.

Background

Learning to read is the most important skill students can learn in school, serving as the very foundation for all other academic subjects (Marchand-Martella, Martella, Modderman, Petersen, & Pan, 2013). Yet the vast majority of upper elementary and middle school students struggle reading grade level text with ease and understanding. The 2013 National Assessment of Educational Progress findings in reading (National Center for Education Statistics [NCES], 2013) showed only 34% of fourth-grade students scored at or above the proficient level; for eighth graders, only 27% of students scored at or above the proficient level. Students scoring at the proficient level "demonstrate solid academic performance and competency over challenging subject matter" (NCES, 2013, p. 7).

To tackle the U.S. literacy challenge to enhance college and career readiness in a comprehensive and cohesive manner, the Common Core State Standards in English Language Arts (CCSS ELA Initiative, 2012) were developed. These Standards provide teachers and parents with a common understanding of what students are expected to learn in reading, writing, language, and speaking and listening to handle 21st century literacy demands. Programs developed to meet the CCSS should be field-tested to ensure their effectiveness with students. This field-testing should occur across multiple sites, teachers, and student populations. Information such as Lexile® growth, number of lessons completed, and teacher feedback can be gathered. This information may be used to make programmatic changes to ensure instructional effectiveness, efficiency, and satisfaction.

Lexile® growth is an important measure of student reading performance. Lexiles® are used to assess the difficulty of a text (i.e., semantic and syntactic elements) and the skill level of a student (MetaMetrics, 2013). They range from BR (Beginning Readers, or individuals who lack decoding strategies to read connected text fluently and who are reading below 0L on the Lexile® scale), to 1700. Lexiles® allow teachers to set goals for individual student growth (Archer, 2010/2011) and determine if students meet yearly growth expectations (Knutson, 2011). These growth expectations were determined using a representative sample of 373,880 students in grades 3 to 10 (Knutson, 2011) and are based in part on students' initial Lexile® scores obtained during the fall benchmark period. Because students initially performing far below grade level find it difficult to catch up and make gains similar to their typically-performing peers, more rigorous yearly growth expectations are required. Based on these growth expectations, the lower the initial Lexile® students receive in the fall, the higher their expected growth scores; the higher the initial Lexile® scores in the fall, the lower their expected growth scores. Thus, it is important to ensure students do not fall behind as it is more difficult to catch up than to make typical, grade-level progress. Such a high stakes strategy argues in favor of interventions to prevent students from falling behind in the first place.

The program evaluation had two purposes. The first was to examine the association between Lexile® growth and instruction using a pre-publication literacy program during field-testing. The second purpose was to determine teacher satisfaction with the program.

Methodology

Participants and Settings

The sample of students in this study was one of convenience. Sixty-nine students in grades 3 to 8 across five states (Washington, Georgia, South Carolina, Tennessee, and Texas) were identified by their respective schools as performing at least 2 years below grade level in reading and thus were selected to receive strategic reading intervention support. Table 1 shows the demographic information for these students across state sites. The majority of these students were Caucasian (about 62%) and male (about 61%). Further, 43% of students were identified with special needs. Approximately, 78% of the students qualified for free and/or reduced price meals. All classroom teachers involved in the evaluation were female.

Table 1: Site Demographic Data

	Site					
	WA	GA	SC	TN	TX	Total
Number of students	9	20	13	16	11	69
Gender						
Male	7	10	8	11	6	42
Female	2	10	5	5	5	27
Grade						
3	0	6	0	0	0	6
4	2	4	0	0	0	6
5	1	10	13	3	0	27
6	6	0	0	2	11	19
7	0	0	0	2	0	2
8	0	0	0	9	0	9
Free or Reduced-Price Lunch	-	19	12	8	10	49
Unknown	9	-	-	-	1	10
Special Education	-	1	11	16	1	29
Unknown	9	-	2	-	1	12
Ethnicity						
Caucasian	-	14	11	16	2	43
African American	-	6	1	0	3	10
Hispanic	-	0	0	0	5	5
Unknown	9	0	1	0	1	11

Washington. The Washington (WA) site was one of 28 elementary schools located within an urban district. It served 610 students in grades K-6. Of the 69 students, 9 (13%) were from this site. Information on free or reduced price status, special education status, and ethnicity was not released by the school. Three teachers implemented the program; years of teaching experience were 1, 7, and 15 years. The program was implemented during an after-school program in the school.

Georgia. The Georgia (GA) site was one of 12 elementary schools contained within a large school district. It served 610 students in grades preschool–5. Of the 69 students, 20 (29%) were from this site. One teacher implemented the program. She had 30 years of teaching experience and taught in a general education classroom.

South Carolina. The South Carolina (SC) site was one of five elementary schools and was considered rural. It served 532 students in grades preschool-5. Of the 69 students, 13 (18.8%) were from this site. One teacher with 8 years of teaching experience implemented the program in a special education resource classroom.

Tennessee. The Tennessee (TN) site was one of four middle schools in a large district. It served 689 students in grades 5-8. Of the 69 students, 16 (23.2%) were from the TN site. Two teachers implemented the program with 11 and 21 years of experience, respectively. Both teachers implemented the program in a special education resource classroom.

Texas. The Texas (TX) site was one of 20 middle schools located within a large district. It served 910 students in grades 5 and 6. Of the 69 students, 11 (15.9%) were from the TX site. The teacher had 15 years of teaching experience. The program was implemented in a general education classroom.

Materials

The schools implemented a pre-publication version of *SRA FLEX Literacy* (Marchand-Martella et al., 2014). This program included three experiences (i.e., digital, print, and project) within each of two systems (elementary and secondary). Sample lessons and program specifics across systems and experiences can be found at http://www.flexliteracy.com/sampler/register and http://flexliteracy.com/reading-intervention-program.

The digital experience provides computer-based learning with vignettes and animated learning opportunities; in this 25-minute, individualized experience, students read and interact with text and earn points toward rewards.

The print experience provides teacher read aloud and shared reading experiences so that struggling students can access more complex, print-based text. For 25 minutes, teachers lead a group of students in discussion and individualized skill application, with in-depth focus on developing deeper comprehension strategies. Individualized teacher-led remediation based on the digital experience, extension activities to enhance critical thinking, and trade books are central to this learning experience.

The project experience helps students build college and career readiness skills in writing- and literacy-centered projects that include research, presentation collaboration, reflection, and evaluation. For 50 minutes, students apply the skills and strategies they learned in the digital and print experiences to science, social studies, and health projects to make cross-curricular connections.

Dependent Variables and Measures

Three primary dependent variables were used in this evaluation: program completion, reading improvement, and program satisfaction by teachers.

Program completion. Program completion was determined by dividing the number of lessons completed by the number of instructional days available for the program (180). To maintain program integrity, the completion of one lesson per instructional day was the goal. One lesson per day was selected to ensure adequate student progress in the program.

Reading improvement. Reading improvement was determined by subtracting each student's pre-program placement score on the Lexile® assessment from the highest benchmark score obtained during the program (as recommended by Archer, 2010/2011). Lexile® assessments are embedded at quarterly benchmark intervals within the digital experience. At the start of the program, students participated in a pre-program placement assessment, the result of which served as the pre-program Lexile® score. This score was subtracted from the highest benchmark score to obtain a Lexile® growth score. Growth scores were compared to the expected growth based on fall scores and needed growth to reach the 50th percentile in Lexile® performance.

Program satisfaction. Participating teachers completed a survey to determine teacher perceptions of the program's impact and to solicit feedback about the program (see Table 2) at the end of the academic year. The teacher survey

included 10 items related to program satisfaction. The survey was constructed solely for this study and was not used in previous studies. Survey items were rated on a Likert-like scale, with response categories ranging from 1 (Strongly Disagree) to 6 (Strongly Agree). Cronbach's alpha for the 10 items was .617.

Table 2: Teacher Satisfaction Survey and Results

Item	Strongly Disagree	Moderately Disagree	Mildly Disagree	Mildly Agree	Moderately Agree	Strongly Agree
SRA FLEX Literacy addresses content standards.	0	0	0	0	5	3
I felt successful teaching *SRA FLEX Literacy*.	0	0	0	1	3	4
SRA FLEX Literacy made my students feel successful.	0	0	0	0	3	5
The students found the lessons engaging.	0	0	0	1	3	4
My students felt successful during the *Print Experience*.	0	0	1	0	2	5
The *Print Experience* was engaging for my students.	0	0	0	1	2	5
My students felt successful during the *Project Experience*.*	0	0	0	2	3	2
The projects were engaging for my students.**	0	0	0	1	3	3
My students felt successful during the *Digital Experience*.	1*	0	2	1	0	4
I would recommend *SRA FLEX Literacy* to others.	1*	0	0	1	2	4

 * Note: Teacher indicated that her responses reflected frustration with technology at this particular site.
** One teacher did not respond to this question.

Procedures

The literacy program was implemented in the fall of the 2012-2013 school year. The approximate dates for program implementation were late September for the SC site, early to mid-October for the WA, GA, and TN sites, and early November for the TX site. Prior to implementation of the program, teachers were provided a half day of on-site training for the print and project experiences, and lessons in the digital experience were demonstrated to the teachers. Five consultants from the program publisher provided the training that included modeling lessons, guiding participants through lessons, and coaching and giving feedback on teacher performance. They served as the primary point of contact, conducted classroom visits, and collected all data. All consultants held a bachelor's (TN) to doctoral (GA) degree in teaching with 7 (TN) to 10 (GA, SC, TX, WA) years of administrative and/or teaching experience.

The time allotted for instruction, degree of program implementation, and the setting in which the program was used varied by site. Two of three teachers at the WA site allotted between 45 to 60 minutes per day for instruction, and one teacher allotted 30 minutes per day for instruction. The GA, SC, TN, and TX sites allotted 90 to 120 minutes per day for instruction. The WA site utilized the print and digital experiences, and the GA, SC, TN, and TX sites implemented the print, digital, and project experiences. All classrooms contained an interactive whiteboard/ Smartboard and computers with two exceptions—the WA and SC students accessed computers in a computer lab rather than the classroom.

Procedural Fidelity

Consultants reported findings via narratives, along with a single final, summative rating on a scale of 1 (extremely low fidelity of program implementation) to 10 (highest degree of program implementation). The summative rating reflected the consultants' holistic observations complied throughout the study. The average ratings ranged from 5 (GA) to 9 (SC). At the GA site, concerns centered on failure to follow all program requirements and deviations from the program script. The SC site classroom was considered a model for implementation by the consultant.

Analysis and Results

Data analysis and interpretation was conducted by an independent evaluator (first author); the evaluator was permitted to publish findings without publisher oversight. Additionally, the program authors (second and third authors of this paper) were not involved in data analysis and interpretation of the results.

Program Completion

The percentage of yearly lessons completed varied significantly by site, $F(4, 64) = 72.83$, $p = .000$. The mean percentage of yearly lessons completed was 29.32% (see Table 3).

Table 3: Lexile® Growth Across Sites

Site	Number of Students	Mean Entry Lesson	Mean Ending Lesson	Mean Number of Lessons	Mean % Complete (1/day for 180 days)	Mean Placement Lexile	Mean Highest Lexile	Mean Lexile Growth
WA	9	139.33	147.44	9.11	5.06%	543.89	600.00	56.11
GA	20	111.25	159.30	49.05	27.25%	401.75	578.25	176.50
SC	13	10.23	130.08	120.85	67.14%	73.46	390.77	317.31
TN	16	74.75	120.50	46.75	25.97%	362.81	515.63	152.82
TX	11	156.45	179.09	23.64	13.13%	577.73	657.73	80.00
Combined (weighted)	69	94.62	146.41	52.78	29.32%	473.75	543.91	166.30

Lexile® Growth

The mean Lexile® growth was 166.30L across the five sites (see Table 3). As shown in Table 4, the mean Lexile® growth ranged from 92.11L (Grade 6) to 216.30L (grade 5). Students in grades 5, 7, and 8 exceeded the expected Lexile® growth. None of the grades overall met or exceeded the needed Lexile® growth to reach the 50th percentile for the spring Lexile® measures.

Table 4: Lexile® Growth Across Grade Levels

Grade	Number of Students	Mean Entry Lesson	Mean Ending Lesson	Mean Number of Lessons	Mean % Completed (1/day for 180 days)	Mean Placement Lexile	Mean Highest Lexile	Mean Lexile Growth	Mean Expected Growth[1]	Mean Needed Growth[1,2]
3	6	54.33	117.33	64.00	35.56%	305.00	466.67	161.67	189.00	285.00
4	6	77.67	124.50	47.83	26.57%	261.67	465.83	204.17	221.00	438.33
5	27	85.26	160.22	75.96	42.20%	315.74	532.04	216.30	196.00	494.26
6	19	131.26	152.79	22.52	12.51%	507.63	599.74	92.11	96.00	372.37
7	2	61.00	102.00	42.00	23.33%	347.50	527.50	180.00	151.00	607.50
8	9	91.00	135.33	45.33	25.18%	420.00	568.89	148.89	139.00	580.00

[1]Based on Growth Expectations by Knutson (2011)
[2]50th Percentile Spring Target Measure

Overall, 33 of the 69 students (47.8%) exceeded the expected Lexile® growth from fall to spring. The numbers of students who exceeded the expected Lexile® growth by grade level were as follows: three of six for grade 3, two of six for grade 4, 14 of 27 for grade 5, nine of 19 for grade 6, one of two for grade 7, and four of nine for grade 8.

Relationship between Lexile® growth and program completion. The correlation between Lexile® growth and percentage of yearly lessons completed for the program was calculated and determined to be statistically significant, $r = .559$, $n = 69$, $p < .001$ (two-tailed). Multiple regression analysis was used to determine the association between Lexile® growth and instruction, after controlling for the number of lessons completed and the program experiences implemented (addition of the project experience). The results of the regression indicated that the number of lessons completed and the program components implemented explained 31% of the variance ($R^2 = .318$, F [2, 66] = 15.381, $p = .000$). It was determined that the number of lessons completed significantly accounted for Lexile® growth, ($b = 1.86$, $t[66] = 4.66$, $p = .000$). The addition of the project experience was not found to be statistically significant, ($b = 32.99$, $t[66] = .706$, $p = .483$).

Program Satisfaction

Overall, teachers generally found the program to meet the CCSS and to be a positive experience for their students (see Table 2). All or the majority of teachers selected either Strongly Agree or Moderately Agree to all questions with the exception of one—students feeling successful during the digital experience. The teachers reported the main issue with this category was frustration with technology issues (these issues were rectified in the published version of the program). Most importantly, six of the eight teachers reported they would recommend the program to others, with one teacher mildly agreeing and one strongly disagreeing. Again, the teacher strongly disagreeing had concerns over technology issues.

Discussion

This program evaluation had two purposes. The first purpose was to examine the Lexile® growth data of students who were exposed to a pre-publication literacy program during field testing. The second purpose was to determine the teacher's satisfaction with the program. Ultimately, the results of this evaluation were used to make programmatic revisions to enhance ease of implementation to better meet the needs and expectations of teachers.

Positive findings were evident in this investigation. First, although the majority of the yearly lessons were not conducted, the mean Lexile® growth across sites and grade levels was statistically significant. In fact, on average, students in grades 5, 7, and 8 exceeded the expected Lexile® growth. This is notable, because once behind it is very difficult for lower-performing students to make the necessary expected gains. Although the average Lexile® growth for students in grades 3, 4, and 6 was below the level expected for a year of instruction, their average gains were less than 28L of what would be expected. Also, almost 48% of students exceeded the expected Lexile® growth. Importantly, there was a statistically significant relationship between Lexile® growth and percentage of yearly lessons completed. Thus, the more lessons students completed, the greater gains they made. This result can be seen when comparing the SC site to the other sites. These results underscore the difficulty students at risk face when they fall behind in their reading skills.

Another positive finding of the evaluation was that the teacher perceptions of the program were positive. Six of the eight teachers reported they would recommend the program to others. The only concern voiced by teachers related to technology issues (e.g., web access, comfort with computers), underscoring the need for teacher training and support with regard to technology.

Limitations

Although positive findings were seen in this evaluation, there were limitations. First, because this was a program evaluation of the pre-publication program, an experimental design was not used to determine cause-and-effect relationships between the program and the outcome effects.

Second, there was a lack of program completion over the period of a year. Due to the nature of the evaluation—to provide ongoing feedback on the strengths and weaknesses of the program for further refinement—there were several "starts and stops" along the way. Therefore, it is not known what the effects of a year-long implementation would be.

Third, there was some variation in implementation across sites. This variation was likely due to the program evaluation aspect of the investigation where there was continual refinement in the program. Therefore, it is not known if, and how, this variation of instructional delivery affected student performance.

Conclusions

Although there were several limitations with this evaluation, results suggest that the targeted literacy program (now revised based on the gathered information) may be an effective one for students who are at risk for reading failure. These field-testing-based results were promising and lay the foundation for further research on the effects of the published version of this literacy program (Marchand-Martella et al., 2014).

References

Archer, L. E. (2010/2011). Lexile reading growth as a function as starting level in at-risk middle school students. *Journal of Adolescent & Adult Literacy, 54,* 281-290.

Common Core State Standards Initiative [CCSS]. (2012). *Common Core State Standards for English language arts & literacy in history/social studies, science, and technical subjects and Appendix A.* Retrieved from http://www.corestandards.org/the-standards and http://www.corestandards.org/assets/Appendix_A.pdf.

Knutson, K. A. (2011). *Growth expectations: Setting achievable goals.* New York, NY: Scholastic.

Marchand-Martella, N. E., Martella, R. C., Fisher, D., McTighe, J., Kosanovich, M., Johnson-Glenberg, M., & Morrell, E. (2014*). SRA FLEX Literacy. Grades 3-12*. Columbus, OH: McGraw-Hill.

Marchand-Martella, N. E., Martella, R. C., Modderman, S. L., Petersen, H. M., & Pan, S. (2013). Key areas of effective adolescent literacy programs. *Education & Treatment of Children, 36,* 161–184.

MetaMetrics. (2013). *Linking assessment with instruction.* Retrieved from http://www.metametricsinc.com

National Center for Education Statistics. (2013). *A first look: 2013 mathematics and reading* [NCES 2014-451]. Washington, DC: U.S. Department of Education, Institute of Education Sciences.

About the Authors

Shannon Flaum-Horvath, Ph.D., works for SKF Educational Services in Ashville, OH. Nancy Marchand-Martella, Ph.D., BCBA-D, and Ronald C. Martella, Ph.D., BCBA-D, are Professors of Special Education at Eastern Washington University in Cheney, WA. Charalambos C. Cleanthous, Ph.D., is an Associate Professor in the Department of Psychology at Eastern Washington University.