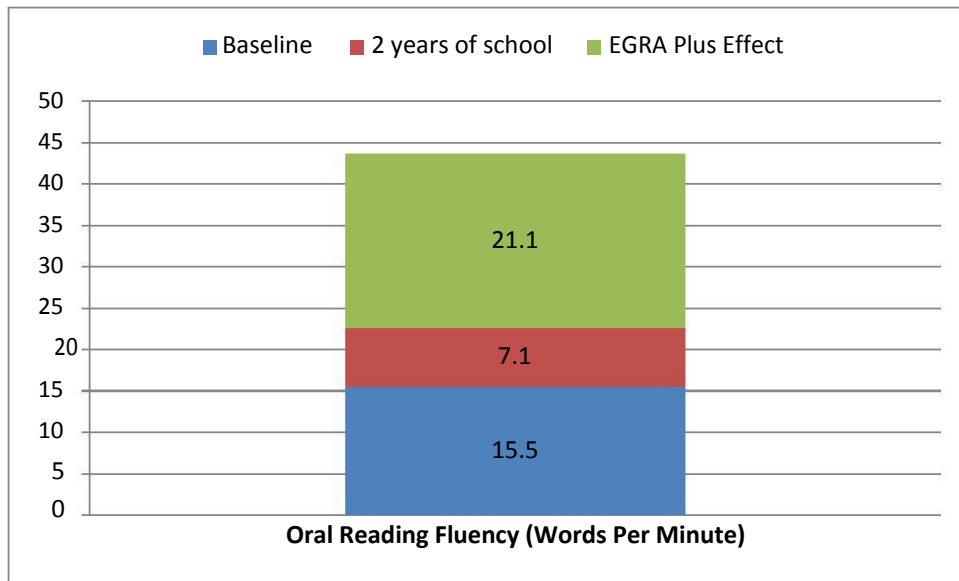




EGRA Plus: Liberia

Program Evaluation Report



Early Grade Reading Assessment (EGRA) Plus: Liberia

EdData II Task Number 6

Contract Number EHC-E-06-04-00004-00

Strategic Objective 3

January 15, 2011

This publication was produced for review by the United States Agency for International Development. It was prepared by RTI International and the Liberian Education Trust.

EGRA Plus: Liberia used The "Fantastic Phonics" decodable reading program

Copyright Jennifer Cooper

All other content copyright to RTI

EGRA Plus: Liberia

Program Evaluation Report

Contract: EHC-E-06-04-00004-00

Prepared for
USAID/Liberia

Prepared by
Benjamin Piper
Medina Korda
RTI International 3040 Cornwallis Road
Post Office Box 12194 Research Triangle Park, NC 27709-2194

RTI International is a trade name of Research Triangle Institute.

The authors' views expressed in this publication do not necessarily reflect the views of the United States Agency for International Development or the United States Government.

1. Executive Summary

Building on the success of the Early Grade Reading Assessment (EGRA) as a measurement tool, many countries have begun to show interest in moving away from assessments alone and toward interventions focused on changing teacher pedagogy, and as a result, increasing student reading achievement. Liberia, for example, began an EGRA-based intervention, called EGRA Plus: Liberia, in 2008. The results from the EGRA Plus midterm evaluation showed very promising results on a variety of learning outcomes.¹ This report is an impact evaluation of the EGRA Plus program at project completion, and it presents compelling evidence that a targeted reading intervention focused on improving the quality of reading instruction in primary schools can have a remarkably large impact on student achievement in a relatively limited amount of time.

Program Design

Liberia's path toward intervention started with a World Bank-funded pilot assessment using EGRA in 2008, which was used as a system-level diagnosis. Based on the pilot results that showed that reading levels of Liberian children are low, the Ministry of Education (MOE) and USAID/Liberia decided to fund a two-year intervention program, EGRA Plus: Liberia, to improve student reading skills by implementing an evidence-based reading instruction program. EGRA Plus: Liberia was designed as a randomized controlled trial. **Three groups of 60 schools were randomly selected into full treatment, light treatment, and control groups.** These groups were clustered within districts, such that several nearby schools were organized together. The intervention was targeted at grades 2 and 3. The design was as follows: The control group did not receive any interventions. In the "full" treatment group, reading levels were assessed; teachers were trained on how to continually assess student performance; **teachers were provided frequent school-based pedagogic support, resource materials, and books;** and, in addition, parents and communities were informed of student performance. In the "light" treatment group, the community was informed about reading achievement using school report cards based on EGRA assessment results or findings and student reading report cards prepared by teachers.

Comparisons at Baseline

Schools in all three groups (control, full treatment, and light treatment) were assessed three times. The baseline measurement took place in November and December 2008,² the

¹ Piper, B., & Korda, M. (2009). *EGRA Plus: Liberia data analytic report: EGRA Plus: Liberia mid-term assessment*. Report prepared under the USAID EdData II project, Task 6, Contract No. EHC-E-06-04-00004-00. Retrieved September 21, 2010, from <https://www.eddataglobal.org/documents/index.cfm?fuseaction=pubDetail&ID=200>

² Baseline: 176 schools were assessed, including 57 control, 59 full treatment, and 60 light treatment schools, for a total of 2,988 students.

midterm in May and June 2009,³ and the final assessment in May and June 2010.⁴

Students were assessed on a variety of essential early grade reading tasks, including letter naming fluency, phonemic awareness, familiar word fluency, unfamiliar word fluency, connected-text oral reading fluency, reading comprehension, and listening comprehension. The tests used for the midterm and final assessments were equated to the baseline assessment in order to ensure comparability of data, and the reliability of the tests was calculated.⁵

Steps also were taken to ensure that the treatment groups were comparable. To illustrate, Table 1 shows the scores for each EGRA section at baseline. Note that this table presents combined scores for grade 2 and 3. The column “Comparison to Control” presents the results of *t*-tests comparing whether the outcome measures for full treatment and light treatment scores were higher or lower than they were for the control schools at the baseline. For full treatment, we found that before the intervention, the full treatment schools had higher average scores on familiar word and unfamiliar word fluency, oral reading fluency, and reading comprehension (at the .10 level). For light treatment, this table shows that light treatment schools outperformed their control school counterparts in oral reading fluency (at .10 level), reading comprehension (at .10 level) and listening comprehension.⁶

Table 1: Comparisons at Baseline and Final—Program Effects and Effect Sizes, by Treatment Group and EGRA Section

Section		Baseline			Final		Program Impact		
		Mean	Standard Deviation (SD)	Comparison to control	Mean	SD	% Increase over Baseline	Program Effect	Effect Size
Letter naming fluency (per minute)	Control	60.67	25.17		82.42	24.37	35.85%		
	Full	62.35	24.86		99.26	24.07	59.20%	14.8***	0.52 SD
	Light	60.37	25.82		88.14	24.49	46.00%	6.0***	0.21 SD
Phonemic awareness (out of 10)	Control	3.41	2.32		4.31	2.87	26.39%		
	Full	3.56	2.26		5.96	2.70	67.42%	1.5***	.55 SD
	Light	3.49	2.30		4.86	2.77	39.26%	0.4**	.18 SD

³ Midterm: 175 schools were assessed, including 56 control, 59 full treatment, and 60 light treatment schools, for a total of 2,805 students.

⁴ Final: 175 schools were assessed, including 58 control, 57 full treatment, and 60 light treatment schools, for a total of 2,688 students.

⁵ Analysis of the assessment tool showed that it was reliable. The Cronbach’s alpha results for baseline, midterm, and final assessments showed reliability of 0.85 or higher, which is quite good. Cronbach’s alpha is a measure of how well a set of variables measure an underlying construct (in this case, early grade reading skill).

⁶ Because of these slight differences at baseline, our program impact analyses account for the differences at baseline in all of the models assessed. This is done in two ways, first, by the random selection of schools, and second, by the differences-in-differences identification strategy “removing” the baseline differences statistically.

Section		Baseline			Final		Program Impact		
		Mean	Standard Deviation (SD)	Comparison to control	Mean	SD	% Increase over Baseline	Program Effect	Effect Size
Familiar word fluency (per minute)	Control	8.51	13.54		18.83	17.41	121.27%		
	Full	10.03	14.28	Higher*	34.88	22.62	247.76%	14.3***	0.78 SD
	Light	9.24	13.86		19.73	20.19	113.53%	0.3	No effect
Unfamiliar word fluency (per minute)	Control	1.91	5.55		2.85	8.73	49.21%		
	Full	2.51	6.22	Higher*	14.70	17.31	485.66%	11.2***	1.23 SD
	Light	2.30	6.22		3.27	7.98	42.17%	1.1*	No effect
Oral reading fluency (per minute)	Control	18.14	19.42		25.21	25.52	38.97%		
	Full	20.83	20.26	Higher**	49.61	33.86	138.17%	21.1***	0.80 SD
	Light	19.77	20.37	Higher~	27.93	29.38	41.27%	1.1	No effect
Reading comprehension (% correct)	Control	23.70	23.86		31.50	33.27	32.91%		
	Full	25.81	24.37	Higher~	59.38	35.49	130.07%	25.2***	0.82 SD
	Light	25.74	24.44	Higher~	34.34	35.61	33.41%	0.7	No effect
Listening comprehension (% correct)	Control	32.64	21.56		69.43	32.33	112.71%		
	Full	33.58	20.11		83.53	24.44	148.75%	13.1***	0.39 SD
	Light	34.51	19.84	Higher*	71.79	31.22	108.03%	0.7	No effect

Legend: *** <.001, **<.01, *<.05, ~<.10

Program Impact

Table 1 also shows the increase in scores over the baseline for each group. Note that these estimates come from a simple tabulation of the data.

- *Letter naming fluency.* At the baseline, Liberian children were capable of identifying the names of letters, with the average control child identifying 60.7 letters in a minute. At the baseline, the letter naming scores were good, which suggests that program impacts were not likely to be very large. At the final assessment, students in full treatment schools showed a 59.2% increase in letters read, while light treatment schools increased in letter fluency by 46.0%. Control schools also increased their scores, by 35.9%. This is evidence of the learning effect, since the final assessment was held at the end of the academic year while the baseline was at the beginning of the year. These were larger impacts on letter naming than we expected, and with respect to program impact, the increases for full treatment were 0.52 standard deviations (SD) and 0.21 SD for light treatment.⁷
- *Phonemic awareness.* Program impact on phonemic awareness was also large. The combined scores for grades 2 and 3 show that the number of sounds

⁷ Note that the effect sizes reported here are Cohen's *d* from the differences-in-differences analyses presented in sections below. Small effect sizes are from 0 to .40, moderate from .40 to .75, and large higher than .75.

identified increased by 67.4% and 39.3% in full and light treatment schools respectively, compared to 26.4% for control schools. This equates to an effect size of 0.55 SD in full treatment schools and 0.18 SD in light treatment schools. **This represented a substantive increase of 2.4 words correct (out of 10) for full intervention schools and 1.4 for light intervention schools.**

- *Familiar word fluency.* **For familiar words, children in full treatment schools increased by 247.8% and light treatment schools by 113.5%.** Since control schools increased their skills as well, the effect size was 0.78 SD for full treatment and was statistically insignificant for light treatment. This is because control-school children increased their scores by 121.2%. This represents an increase of 24.9 and 10.5 words per minute.
- *Unfamiliar word fluency.* For unfamiliar words, control and light treatment schools had very limited changes in outcomes. Control schools increased by 0.9 words (49.2%), while light treatment schools increased by 1.0 word per minute (42.2%). **For full treatment schools the increase was 485.7% (from 2.5 words per minute to 17.3 words per minute).** The effect size was a very large 1.23 SD for full treatment and insignificant for light treatment.
- *Oral reading fluency.* The impact was also quite large for fluency in oral reading of connected text. **Compared against baseline, full treatment children increased the number of words read correctly by 138.2%,** light treatment schools increased by 41.3%, and control schools by 39.0%. Substantively, this means that full treatment schools increased their number of words read from 20.8 to 49.6 words per minute, while light treatment increased from 19.8 to 27.9. Compared against the gains for control schools, these effect sizes are positive for full treatment (at 0.80 SD) and insignificant for light treatment. **This means that at the final assessment, children in full treatment schools were reading nearly two and a half times as fluently as they were at the baseline.**
- *Reading comprehension.* Comparing the final and baseline assessment scores in reading comprehension, we find that **full treatment schools increased their scores by 130.1% over baseline, while light treatment scores increased by 33.4% and control schools by 32.9%.** This means that, at the final assessment, children in full treatment schools scored 33.6 percentage points higher than control-school children scored at the baseline, with students in light treatment schools scoring 8.8 percentage points higher. The program's effect size was 0.82 SD and the effect was statistically significant for full, but insignificant for light treatment schools. **This is more than a doubling of the reading comprehension percentage rates for full treatment children.**
- *Listening comprehension.* For listening comprehension, the increases for full and light treatment schools were 148.8% and 108.0%, respectively. It should be noted that control schools increased their scores by 112.7%, so only by taking into account the baseline scores can a true program effect be estimated. Substantively,

full treatment schools increased by 49.9% and light treatment schools by 37.3% over baseline, an effect size of 0.39 SD and no effect, respectively.

Comparing across EGRA sections, we find that the EGRA Plus full treatment program had moderate impacts on listening comprehension, large impacts on phonemic awareness, letter fluency, familiar word fluency, oral reading fluency, and reading comprehension. We found very large impacts for unfamiliar word fluency, indicating that the EGRA Plus program had particularly large impacts on improving children's ability to manipulate sounds to make words.

Sex and Grade Differences

On all EGRA sections, grade 3 students scored statistically significantly higher than grade 2 students, with more than 11 additional words read correctly per minute on the oral reading fluency section. This is a measure of standard intergrade improvement, which we note for the sake of comparison: The project impact was much bigger than the standard intergrade improvement. In other words, the project was able to boost children's learning by much more than one grade. On the other hand, there were no differences between boys' and girls' achievement, except for unfamiliar word fluency, where girls outperformed boys. This appears to be largely because the EGRA Plus program had a slightly larger impact for girls than for boys, partly because scores were lower for girls at the baseline. This suggests that more work is necessary so that the program, when it is folded into other efforts in Liberia, increases the skills of boys in the more complex portions of reading, and careful attention must be given to ways to ensure that boys and girls benefit equally from the program.

Overall Program Impact

This report presents the effect sizes from a more sophisticated analysis using differences-in-differences analyses.⁸ These are presented in Table 1 above, in the effect size column. These analyses show that the full treatment group increased student achievement for every section of the EGRA, often with quite large impacts on student achievement. In fact, the overall EGRA Plus effect size was 0.79 standard deviations, which is enormous in social science. When the program impacts are expressed in terms of grade effects, the full treatment increased letter naming fluency by 1.2 times the effect of being in school for one year. Amazingly, this was the *smallest* effect size for any of the skills assessed. The EGRA Plus full treatment effect was the equivalent of 1.9 school years in phonemic awareness, 1.8 school years in familiar word reading, a remarkable 8.0 years in

⁸ Differences-in-differences is an identification strategy that attempts to make causal inference about a treatment effect by removing the secular trend using a pre and post treatment-and-control design. It is preferable to use three waves of data, if possible, as this particular data set allows. See Skoufias, E., & Shapiro, J. (2006). *The pitfalls of evaluating a school grants program using non-experimental data*. Working paper. Washington, DC: World Bank. Retrieved September 30, 2010, from http://www.webmeets.com/files/papers/LACEA-LAMES/2006/390/pec_eval.pdf

unfamiliar word fluency,⁹ 1.9 years in oral reading fluency, 2.0 years in reading comprehension, and 1.8 years in listening comprehension. The light treatment group also increased student achievement in letter fluency and phonemic awareness.

This report shows that full treatment schools dramatically accelerated children's rates of learning. Our regression estimates show that full treatment children increased their word naming fluency by 2.1 words per minute per month, while the associated rate for control schools was an increase of 0.8 words per minute per month. The rate, then, in full treatment schools, was 2.6 times as fast. For unfamiliar word fluency, we find that the increase in fluency scores was 12.4 times faster in full treatment schools than in control schools, which suggests that a primary entry point for improving reading outcomes for students was through improved decoding skills. The relationship between full treatment and control schools for oral reading fluency of connected text was 4.1 times faster, and 4.0 times faster for reading comprehension. **This shows that the EGRA Plus program did not simply increase the learning outcomes for children; it dramatically accelerated children's learning to an extent seldom found in educational or social science research.**

In summary, given the existing literature on effect sizes in literacy interventions, **EGRA Plus: Liberia far exceeded expectations with respect to impact on student achievement, particularly in the full treatment schools.** Note that the effects were most often large in full treatment schools, with some moderate effect sizes. The range of effect sizes for full treatment was from 0.39 to 1.23 SD. **Oral reading fluency scores at the final assessment were 238% of what they were at the baseline, reading comprehension scores were 230% of what they were at baseline, and decoding fluency scores were 585% what they were at the baseline assessment. These impressive results were found on all of the essential early reading components. Most critically, the EGRA Plus program accelerated the learning of children so much that children learned the equivalent of *three years* of schooling in one year.**

Program Impact Compared with Expectations

When compared against the Performance Management Plan (PMP) of February 2009, the results from the EGRA Plus program are very strong. The PMP noted that the impact over baseline two years later would be a 35% increase for oral reading fluency and reading comprehension in full treatment schools, while light treatment schools would see a 10% increase for those same tasks. For both boys and girls, for both grade 2 and grade 3, the light treatment schools made their target in oral reading fluency. By the same token, for both sexes and both grades, the full treatment schools increased by more than 30%, and for each disaggregated level, the increase was more than 100%. The results

⁹ Note that the comparison is between the effect of moving from grade 2 to grade 3. For unfamiliar words, one must assume that the rate of learning to decode will increase as children get older; that is, the program effect is not linear. That said, children in full treatment schools benefited a significant amount in this section.

were similar for reading comprehension. The increases for both boys and girls in grades 2 and 3 in light treatment schools were more than 10%, and the impacts for all groups in full treatment schools were over 80%.¹⁰

Table 2: Disaggregated Analysis of Percentage Increases Over Baseline, by Treatment Status, Grade, and Sex

Section	Treatment	Grade 2		Grade 3	
		Boys	Girls	Boys	Girls
Oral reading fluency	Control	4.63%	86.05%	32.71%	55.40%
	Full	152.73%	244.36%	89.01%	129.36%
	Light	33.77%	71.63%	21.35%	18.81%
Reading comprehension	Control	4.82%	51.17%	32.04%	46.93%
	Full	149.16%	192.49%	84.76%	209.16%
	Light	17.67%	49.18%	18.81%	53.91%

EGRA Plus Increased Mathematics Outcomes

Interestingly, the data show that while there was no intervention in any subject other than reading, the **EGRA Plus program is likely to have increased mathematics outcomes. This was assessed by comparing the mathematics scores for children in different treatment groups. We found that for full treatment, EGRA Plus increased math scores in number identification, quantity discrimination, addition, subtraction, multiplication and fraction knowledge.** For light treatment, scores increased for multiplication and fractions, but decreased in number identification. More research is necessary to determine whether the full treatment effects were due to the close relationship between reading skills and outcomes in other subjects, or whether the pedagogical techniques that the teachers obtained in EGRA Plus were also effective in other subjects. However, it does buttress the point of view that reading can be a starting place for quality improvements in other subjects, and even at higher levels in the education sector.

Recommendations

Given the success of the EGRA Plus program, we make the following recommendations:

- **Scale up the EGRA Plus program.** Given the remarkable success of EGRA Plus, there appears to be an opportunity for the Liberian Ministry of Education to scale up and expand the intervention. The Liberia Teacher Training Program (LTTP2) is a potential incubator for further interventions and offers an opportunity to determine whether the remarkable impacts of this program can be

¹⁰ Note that while these increases were quite high, the scores for control schools also increased, and at nearly the same rate as those of light treatment schools. This is why many scholars prefer reporting the impact of a program over the baseline and over control schools, to remove the “secular trend.”

replicated at scale. For the last calendar quarter of 2010, at USAID’s direction and with remaining EdData II Task 6 funds, RTI expanded the EGRA Plus: Liberia intervention to all schools—control, light, and full—for another semester.

- **Move past focus on letters and words and focus on reading comprehension.** It appears that improving the oral reading fluency and decoding skills of children is quite possible, and this is highly correlated with reading comprehension, as the program results show. However, the effect on comprehension is not as large as it might have been if more emphasis had been placed on developing the metacognitive skills that children need in order to synthesize and understand written text. We know that children can understand a higher percentage of what they hear than what they read. Explicit instruction and modeling is necessary to match children’s listening comprehension with their comprehension after they read written texts.
- **Develop benchmarks for reading.** The wealth of data obtained in the three waves of assessment from EGRA Plus provide enough evidence for the Liberian Ministry of Education to determine what rates of fluency, comprehension, and word skills are necessary at each level. Such a benchmark development process will help to target resources and efforts, to invigorate the efforts to improve educational outcomes.
- **Target reading techniques using professional development.** Liberian teachers have been proven to be receptive to new pedagogical techniques and strategies. With targeted efforts, teachers can improve how well children read, quite quickly. We recommend that the evidence from this program be included in pre-service and in-service teacher professional development programs of the Liberian Ministry of Education going forward. This will require adaptation efforts, to transfer the mechanisms that were so effective in EGRA Plus to the pre-service sector.
- **Improve girls’ reading achievement.** The findings here showed that while boys outperformed girls at the baseline, with instruction and investment, girls could narrow and even close the sex gap. Therefore, education officials can and should demand high achievement for girls in the classrooms under their jurisdiction, and efforts should be made to encourage teachers to have high expectations for girls.
- **Decoding skills must be emphasized.** The largest impacts of EGRA Plus were found in the tasks that measured children’s ability to decode and to use the alphabetic principle. These skills were mostly lacking in nonproject schools, and it seems that these skills were crucial gateways to the rapid acceleration of learning outcomes that EGRA Plus caused.
- **Use reading improvements to increase learning in other subjects.** The findings showed that reading improvements have the potential for carryover effects in other subjects, in this case mathematics. This suggests that reading is a ripe

subject for interventions, since other subjects might be improved by the simple method of increasing reading outcomes.

- **Expand the use of scripted programs for lesson delivery.** The experience of EGRA Plus makes clear that scripted lesson plans can be a part of an effective program for reading improvement. The increased rates of learning between the midterm and final assessment show that while there was some initial resistance to such methods, the creation of and support for lesson plans for teachers has a high likelihood of continuing to be effective in Liberia.
- **Plan for unmet demand for high-quality education.** The qualitative follow-up report revealed that EGRA Plus not only increased student achievement, but also increased school enrollments. This suggests that there is unmet demand for high-quality education, and that improving reading outcomes will encourage student retention. Enrollment increases in full treatment schools were more than 30%, which has implications for teacher deployment and planning projections throughout the school system.

2. Introduction

The Early Grade Reading Assessment (EGRA) Plus: Liberia program (2008–2010) was an experimental intervention. The intervention was part of a joint collaboration among the Liberian Ministry of Education, World Bank Liberia, and USAID/Liberia. Baseline, midterm, and final assessments were conducted and the results were judged against agreed-upon targets for improved student performance. The baseline assessment was conducted in November 2008, the midterm assessment was conducted in June 2009, and the final assessment took place in June 2010.

The EGRA Plus: Liberia program used empirical data from reading assessments in grades 2 and 3 to track progress toward quality improvements in early grade reading instruction. The research and intervention design allowed for the comparison of three different groups. The first was a control group that received no program interventions, but whose performance was measured (without alerting them to the fact there would be repeated measurement). The second group, the “light” intervention, was a set of schools where parents and community members were provided student achievement data in the area of literacy; they were made aware that there would be testing again. In addition, light intervention teachers were trained in the development of a student reading report card, which they issued four times a year. The final group, the “full” intervention, provided an intensive teacher-training program targeting reading instructional strategies, in addition to the same type of information on student achievement that was provided to parents and communities in light treatment schools. Note that the assignment of schools into treatment groups was random, accounting for geographic clustering.