India Literacy Program

2017 Endline Evaluation Report

Pinaki Joddar
Peter Cooper

August 2017

Room to Read's Literacy Program is a school-based intervention that seeks to develop children's reading skills and reading habits in the early grades. The Literacy Program team in India is working with government counterparts to design an explicit reading instruction program. The program intends to provide a strong foundation in reading and writing skills for all children in Room-to-Read-supported schools, with the goal that children will become fluent readers by the end of Grade 2. In August 2015, Room to Read began a two-year study to measure the impact of the program on children's reading skills. For the endline data collection that occurred from February to April 2017, Room to Read contracted an external research agency to administer reading assessments to Grade 2 children at 75 schools targeted by the program and 75 comparable schools not targeted by the program. These schools were dispersed evenly across Chhattisgarh, Rajasthan and Uttarakhand. Endline results indicate that the Literacy Program is having a large positive impact on reading skills. Moreover, children in program schools experienced gains in learning that were two to three times larger than those experienced by comparison school children from the beginning of Grade 1 to the end of Grade 2. A key concern for program will be to reach those children who are still performing at low levels.

This page is left blank intentionally

Table of Contents

1 Executive Summary 1
2 Introduction 3
2.1 Literacy Program 3
2.2 Impact of the Literacy Program 3
3 Results 5
3.1 School and Child Background Characteristics 5
3.2 End of Grade 2 Reading Results 5
3.2.1 Changes in Average Reading Scores 5
3.2.2 Fluency and Comprehension Benchmarks 7
3.2.3 Zero Score Prevalence 9
3.3 Comparisons by Gender 9
3.4 Comparison across States 11
3.5 Characteristics of Struggling Readers 14
4 Context and Limitations 16
4.1 Context 16
4.2 Limitations 16
5 Conclusion 17
6 Next Steps 17
7 References 19
Appendix A: Literacy Program Overview 20
Appendix B: Research Design 21
Methodology 21
Sampling 21
Reading Assessments 22
Assessor Training 22
Data Collection 22
Data Entry 23
Data Analysis 23
Appendix C: School and Child Background Characteristics 25
Appendix D: Reading Assessment Results by State 27

1 Executive Summary

Room to Read's Literacy Program is a school-based intervention that seeks to develop children's reading skills and reading habits in the primary grades. The program includes two main components: (1) reading and writing instruction for children in Grades 1 and 2; and (2) access to reading materials through the establishment of school libraries. For the instruction component, the Literacy Program team in India worked with the government counterparts to design an explicit reading instruction program. The program intends to provide a strong foundation in reading and writing skills for all children in the Room to Read schools, with the goal that children will become fluent readers by the end of Grade 2. The program includes detailed lesson plans, classroom materials, and comprehensive teacher professional development including continuous monitoring and support from literacy coaches.

In India, the Literacy Program currently operates in 1,536 schools across six states; out of which 610 schools in Chhattisgarh, Rajasthan, and Uttarakhand are implementing the comprehensive instruction plus library approach. Implementation of the Literacy Program in these states are receiving financial support from various sources, including United States Agency for International Development (USAID) and Tata Trust.

In 2015, Room to Read began a two-year impact evaluation to determine the impact of the Literacy Program on children's reading skills. The evaluation included 75 schools benefiting from the program and 75 comparable schools not benefiting from the program (comparison group). For both groups, schools were evenly dispersed across Chhattisgarh, Rajasthan, and Uttarakhand. Room to Read hired an external research agency to collect data on children's reading skills using a version of the Early Grade Reading Assessment (EGRA)[1] that was adapted to Hindi by local experts.

Baseline data were collected in August 2015 and indicated that children from the project group entered Grade 1 with the similarly low reading skills as children from the comparison group. On average, children from both groups could read less than one word per minute. In February-March 2016, midline data were collected from the same cohort of Grade 1 children that were assessed at baseline to determine the impact of the program after one academic year. Data revealed that the program had a large positive impact on the development of children's reading skills. By the end of Grade 1, children from program schools could, on average, read nine words per minute on a test of oral reading fluency, while children from comparison schools could read only two words per minute. In February-April 2017, endline data were collected from the same cohort of children who were assessed at both baseline and midline to determine the impact of the program after two academic years.[2] The endline data showed that the large positive gains that were seen after one year of the program continued after two years. By the end of Grade 2, children from program schools could read an average of 32 words per minute, while children from comparison schools could read only 12 words per minute. On a test of reading comprehension, program school children could correctly answer an average of one more question (out of five) than comparison school children. Gains in learning for both in project schools were two to three times more than their counterparts in comparison schools from baseline to the end of Grade 2. Benefits of the program for girls versus boys varied by assessment task. Though the program was effective across all three states, project school children from Chhattisgarh

[1] The EGRA was developed by RTI International in 2006. For more information, please see: https://www.eddataglobal.org/reading/index.cfm?fuseaction=pubDetail&id=929.

[2] Individual students participated in the baseline assessment were not traced during mind-line and endline surveys; rather, we resampled from the same cohort of children at min-line and endline from the same schools.

performed the best and project school children from Rajasthan experienced the largest two-year gains relative to comparison school children.

Overall, the evaluation results validate Room to Read's approach of providing explicit reading instruction, related materials, on-site monitoring and support, and comprehensive professional development to teachers on literacy instruction. However, the results also point to areas for improvement. For example, though the relative impact of the program was large, the overall reading fluency levels of project school children (32 words per minute, on average) was below our goal of at least 45 words per minute by the end of Grade 2. Moreover, 9 percent of project school children (and 43 percent of comparison school children) were effectively non-readers, while 21 percent of project school children (and 58 percent of comparison school children) were unable to answer even one comprehension question correctly.

To improve upon the evaluation results, Room to Read will continue to refine the Hindi Literacy Program, with a focus on:

- Include more oral language activities in the foundation phase of the program.
- Provide more opportunities for teachers to focus on continuous remediation and revision. Specifically, teachers will have space to conduct remedial activities on the sixth day of every week for children who are struggling.
- Provide more opportunities for children to practice comprehension-based texts. This includes making independent reading time part of the regular instruction routine. During the independent reading time, we will create and encourage teachers to use story cards (i.e., simple texts with colorful illustrations) to strengthen the reading experiences of children.
- Provide greater emphasis on analysis and use of formative assessment data in Literacy Coach and teacher trainings.

Through these strategies, Room to Read hopes to achieve its goal of having all children in the program reading fluently by the end of Grade 2.

2 Introduction

2.1 Literacy Program

Room to Read's Literacy Program is a school-based intervention that seeks to develop children's reading skills and reading habits in the primary grades. The program includes two main components: (1) reading and writing instruction for children in Grades 1 and 2; and (2) access to reading materials through the establishment of school libraries. For the instruction component, the Literacy Program team in India worked with the government counterparts to design an explicit reading instruction program. This intends to provide a strong foundation in reading and writing skills for all children in the Room to Read schools, with the goal that children will become fluent readers by the end of Grade 2. The program includes detailed lesson plans, classroom materials, and comprehensive teacher professional development including continuous monitoring and support from literacy coaches (see *Appendix A* for a more detailed description of the instruction component of the Literacy Program). To date, the Literacy Program has had a substantial impact on reading skills in nearly all of the countries in which it operates, including India.

In India, the Literacy Program currently operates in 1,536 schools across six states; out of which 610 schools in Chhattisgarh, Uttarakhand, and Rajasthan are implementing the comprehensive instruction plus library approach. Operation of the Literacy Program in these states are receiving financial support from various sources, though two sources deserve separate mention here:

- In 40 schools in the Sirohi district of Rajasthan, the intervention is being funded by United States Agency for International Development (USAID) and Tata Trust through an agreement with the Center for micro Finance (CmF) under the Nurturing Early Literacy Program (NELP) since 2015. A total of 1,085 children were enrolled in Grade 1 across these 40 schools at the beginning of 2015-16 academic year. In the 2016-17 academic session, another 60 schools were taken up for implementation in Sirohi under this same project.
- In September 2015, Room to Read entered an agreement with the USAID for implementing a large-scale and innovative early grade reading program aimed at benefiting children in the government primary schools in the states of Chhattisgarh and Uttarakhand. The project includes a phase in which Room to Read demonstrates an effective model towards improving reading outcomes among primary grade children through implementation of its Literacy Program. As part of the "demonstration" phase of this agreement, Room to Read has been working in 80 schools in Chhattisgarh and 100 schools in Uttarakhand since the beginning of the 2015-16 academic year.[3]

2.2 Impact of the Literacy Program

To determine the level of impact of the intervention on children's learning, Room to Read conducts evaluations with children in schools benefitting from Literacy Program (project schools) and schools not benefitting from the Literacy Program (comparison schools). In general, data collection for the evaluation occurs at the beginning of

[3] Known as the Scaling up Early Reading Intervention (SERI), this USAID funded project aims to demonstrate an effective model towards improving reading outcomes among primary grade children and an innovative approach for scaling up NGO-led interventions through the government system. The SERI project is based on a scaffold approach which includes: (i) demonstration of the model in government schools by Room to Read (also called demonstration phase); (ii) close collaboration with the government to scale up the model across entire districts; and (iii) handing over the model to the government for replication.

Grade 1 (baseline), the end of Grade 1 (midline), and the end of Grade 2 (endline). Data from the evaluation enable Room to Read to:

1. Determine whether the program is having an impact on students' reading skills after one and two academic years; and
2. Determine whether the implementation of the program facilitates the acquisition of early reading skills in children at a rate that ensures that they will reach the goal of becoming fluent readers by the end of Grade 2; and
3. Identify reading skills that could be better supported by the program and determine how to improve these reading skills quickly and effectively.

In August 2015, Room to Read hired an external research agency to conduct baseline assessments with Grade 1 children in 75 project schools and 75 comparison schools spread evenly across Chhattisgarh, Rajasthan, and Uttarakhand. The baseline results indicated that children from the project group entered Grade 1 with the similarly low reading skills as children from the comparison group. On average, children from both groups could read less than one word per minute. In February-March 2016, midline data were collected from the same cohort of Grade 1 children that were assessed at baseline to determine the impact of the program after one academic year. Data revealed that the program had a large positive impact on the development of children's reading skills. By the end of Grade 1, children from program schools could, on average, read nine words per minute on a test of oral reading fluency, while children from comparison schools could read only two words per minute. In February-April 2017, endline data were collected from the same cohort of children who were assessed at both baseline and midline to determine the impact of the program after two academic years.[4] The results from endline assessments, as well those from the baseline and midline assessments, appear in *Section 3: Results*.

(See *Appendix B* for a full description of the research design, including details around methodology, sampling, reading assessments, training assessors, data collection and entry, and data analysis.)

[4] Individual students participated in the baseline assessment were not traced during mind-line and endline surveys; rather, we resampled from the same cohort of children at midline and endline from the same schools.

3 Results

3.1 School and Child Background Characteristics

Because the intervention was not allocated randomly to project and comparison schools, it is important to assess whether the two groups are comparable. We did this by examining school and child background data together for the three states (Chhattisgarh, Rajasthan and Uttarakhand) included in the endline survey. (See Tables C.1 and C.2 in *Appendix C,* for more details.)

Across all three states, we examined potential differences between project and comparison schools for different background indicators like location (rural and urban), Grade 2 enrolment, type of classroom (mono- and multi-grade), number of in-position teachers, presence of dedicated language-arts teacher, duration of reading instruction, and attendance on the day of the test, respectively. We found that the differences between the two groups across these indicators were not statistically significant.

In addition to school level information, data were collected for each child on age, sex, language spoke at home, whether attended pre-school, whether having television and collection of books at home, and whether receiving tutoring at home, respectively. The only student-level background indicator for which statistically significant differences between project and comparison school children were found was oral language spoken at home among students assessed at endline ($p < 0.001$). Statistical comparisons of learning assessment results between project and comparison groups took this difference into account.

3.2 End of Grade 2 Reading Results

3.2.1 Changes in Average Reading Scores

Table 3.1 provides an overview of the baseline (July-August 2015), midline (February-March 2016), and endline (February-April 2017) assessment results by project and comparison group across each assessment task. **Overall, children from project schools performed better and experienced significantly greater gains than children from comparison schools.** By the end of Grade 2, children from project schools could correctly read an average of 32 words per minute, while children from comparison schools could read just 12 words per minute. For reading comprehension, children from project schools could correctly answer an average of 2.3 questions (out of 5), compared to just one question answered correctly by comparison school children. Additionally, children in project schools made multifold improvement from baseline to endline on all four reading tasks when compared to children in comparison schools. On average, improvements amongst project school children after two-years of intervention were two to three times greater than progress made by their counterparts in comparison-schools.

TABLE 3.1: Reading Assessment Results – Project vs. Comparison

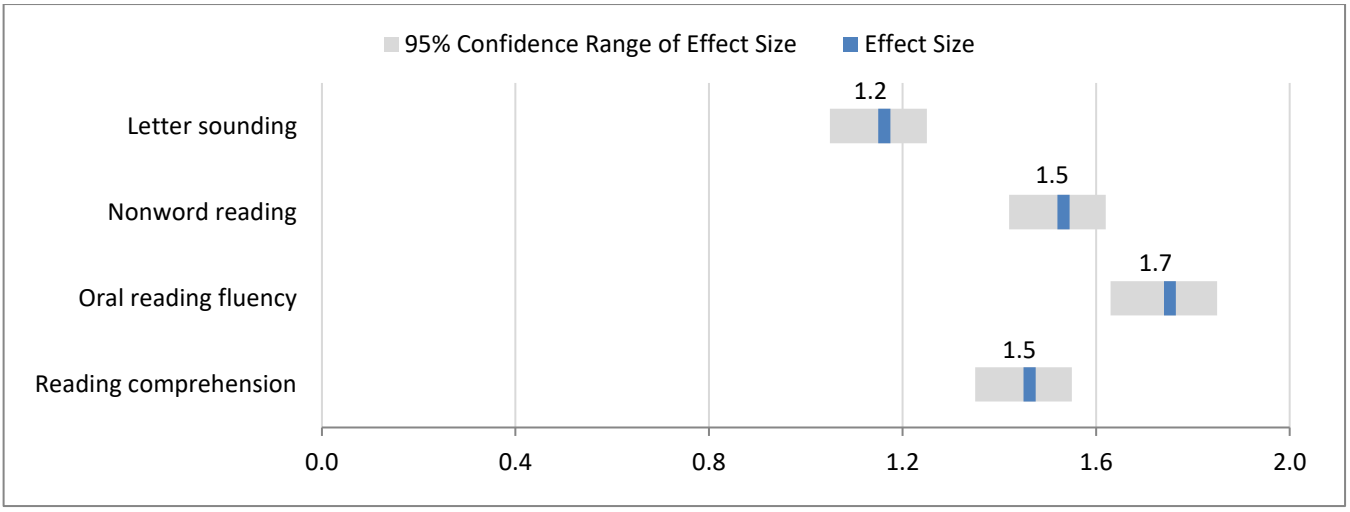
Assessment Task	Group	BASELINE n	BASELINE Mean	BASELINE SD	MIDLINE (END OF GRADE 1) n	MIDLINE (END OF GRADE 1) Mean	MIDLINE (END OF GRADE 1) SD	ENDLINE (END OF GRADE 2) n	ENDLINE (END OF GRADE 2) Mean	ENDLINE (END OF GRADE 2) SD	2-Year Gains	Adjusted Difference in 2-Year Gains over Comparison[†]
Letter sounding fluency	Project	982	3.0	6.6	919	28.6	19.2	933	51.5	24.8	+48.5	+20.6***
(letters per minute)	Comparison	893	2.0	0.6	841	15.3	16.8	882	29.9	22.8	+27.9	
Non-word reading	Project	982	0.3	1.4	919	6.7	7.6	933	16.2	10.8	+15.8	+8.7***
(non-words per minute)	Comparison	893	0.2	1.5	841	2.6	4.9	882	7.4	8.9	+7.2	
Oral reading fluency	Project	982	0.4	2.4	919	8.5	12.0	933	32.2	25.1	+31.8	+20.5***
(words per minute)	Comparison	893	0.2	1.2	841	2.2	5.1	882	11.6	16.9	+11.4	
Reading comprehension	Project	982	0.0	0.2	919	1.0	1.3	933	2.3	1.7	+2.3	+1.3***
(questions answered correctly)	Comparison	893	0.0	0.1	841	0.3	0.8	882	1.0	1.4	+1.0	

Legend of statistical significance of differences between project and comparison schools: *** < 0.001, ** < 0.01, * < 0.05.

[†] Adjusted difference in gains reports the coefficient of the regression analysis conducted with random effects at the school level and age, sex and language spoken at home as covariates. Additionally, version of passage used during the tests were included in the regression analysis for reporting results for oral reading fluency and reading comprehension, respectively.

Differences in child gains are further examined by looking at the adjusted effect sizes for the program across the assessment tasks (See Figure 3.1). The effect size statistic is used to make comparisons across measures that use different scales or units. For the purposes of this analysis, we used the standardized mean effect size statistic, through which an effect size of 0.80 or higher is considered large. We calculated effect sizes by first determining the adjusted difference in gains between project school children and comparison school children through linear regression analysis that took into account the clustering effect (see *Appendix B: Research Design*) and then dividing this difference by the adjusted pooled standard deviation of children' scores. **The effect sizes for Room to Read's Literacy Program were large across all of the tasks, with the largest effect size of 1.7 estimated for oral reading fluency.**

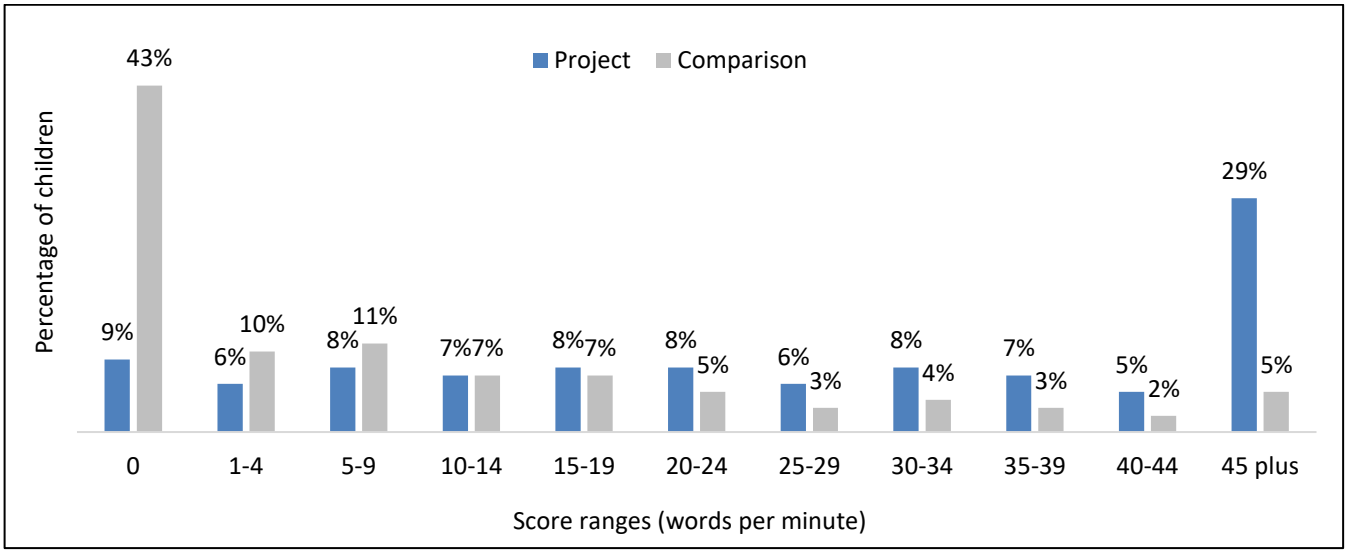
FIGURE 3.1: Adjusted Effect Sizes across Assessment Tasks



3.2.2 Fluency and Comprehension Benchmarks

Although there has been limited research into fluency in Hindi language, independent studies in multiple countries have shown that children at the end of Grade 2 need to reach a fluency rate of roughly 45 to 60 words read per minute (or local-language equivalent) as a prerequisite for reading with comprehension (Abadzi, 2011). We aim for children to reach this fluency level by the end of Grade 2 in India. The distribution of oral reading fluency presented in Figure 3.2 is indicative of how we have progressed to achieve the goal by the end of Grade 2. **Approximately 29 percent of children from project schools met or exceeded the fluency benchmark of 45 words per minute by the end of Grade 2 (compared to only 5 percent comparison school children).** Moreover, differences between the oral fluency scores of children from project and comparison schools were statistically significant ($p < 0.001$).

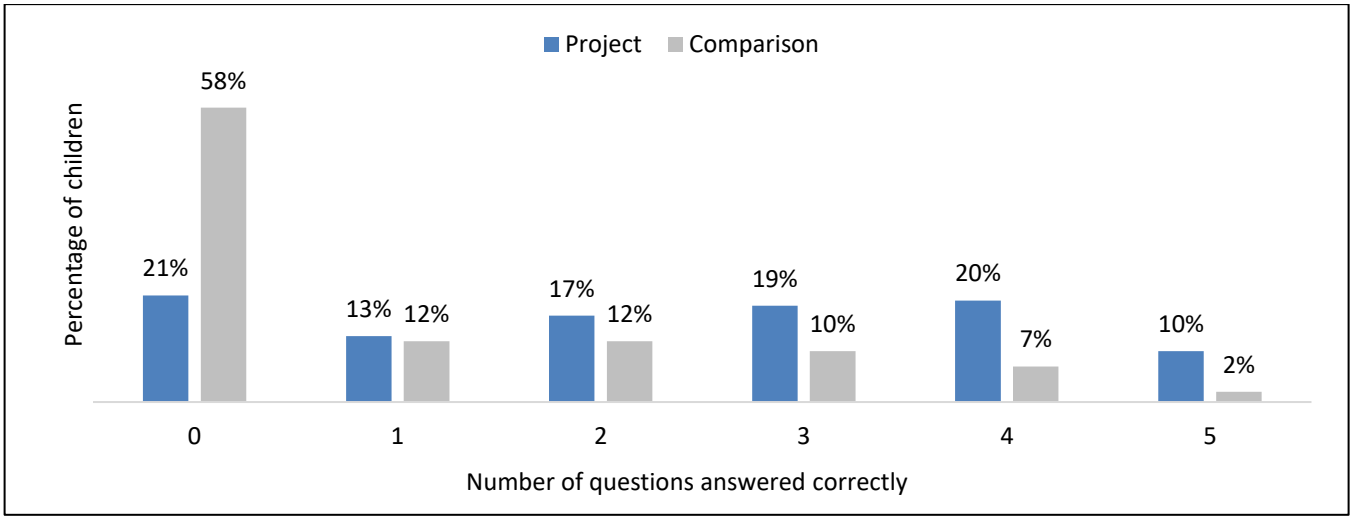
FIGURE 3.2: Distribution of Oral Reading Fluency Scores at Endline



Emergent readers, defined by children who can read 20 or more words per minute, is another internal indicator for oral reading fluency of children. **The endline data revealed that 63 percent of the children from the project schools could read at least 20 words per minute as compared to just 22 percent for comparison school children.** This difference was statistically significant ($p < 0.001$).

Other benchmark indicators that reflect progress include the proportion of children answering at least sixty percent (3 out of 5) and eighty percent (4 out of 5) questions correctly on the reading comprehension task. As shown in Figure 3.3, **49 percent of the project school children could correctly answer at least sixty percent of the comprehension questions correctly (compared to just 19 percent comparison school children), and 30 percent of the project school children could correctly answer at least eighty percent of the comprehension questions correctly (compared to just 9 percent comparison school children).** The difference between project school children and comparison school children on each of these benchmarks was statistically significant (p <0.001).

FIGURE 3.3: Distribution of Reading Comprehension Scores at Endline

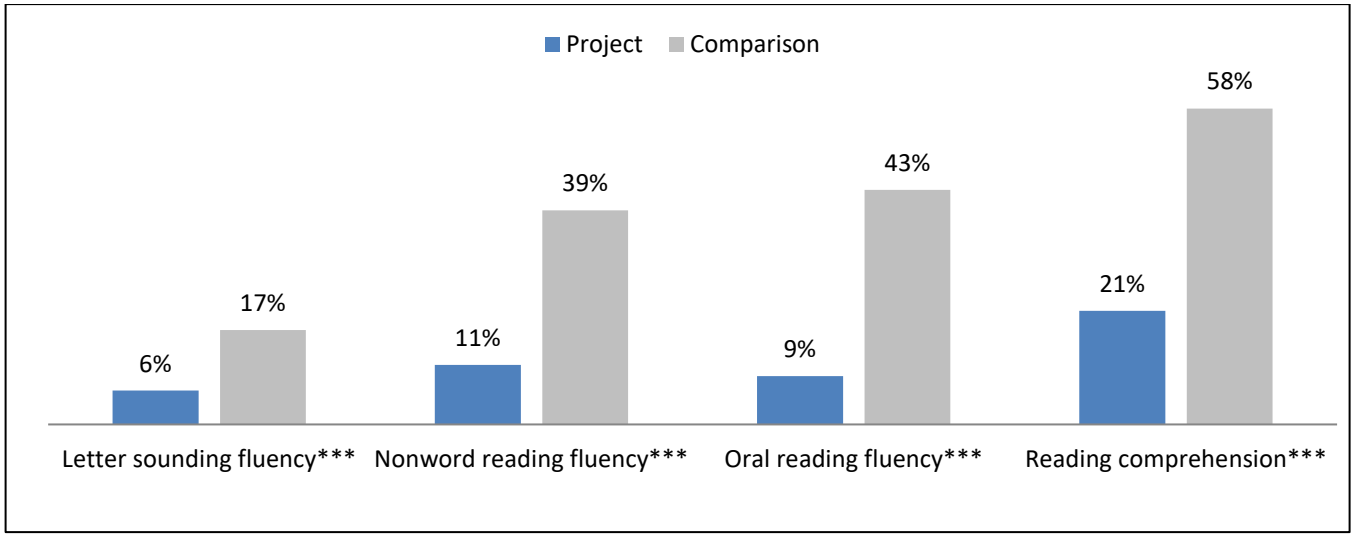


3.2.3 Zero Score Prevalence

Analysis of zero scores provides another view of how children performed, with a focus on children with the lowest achievement. In the Early Grade Reading Assessment, zero scores include those instances in which a child does not provide correct responses to any of the items in a particular assessment task, as well as those instances when a child does not answer or respond correctly to any item in the first line of the assessment task (also known as a discontinued task). Zero scores on tasks show the subset of children who can be characterized as nonreaders.

Figure 3.4 below compares zero scores between the project and comparison school children. **Percentages of zero scores were lower in project schools than comparison schools across all four assessment tasks**, and these differences were statistically significant ($p < 0.001$). In project schools, the highest proportion of zero scores were noted for reading comprehension (21 percent), followed by non-word reading fluency (11 percent). Overall, these results suggest that some project school children still struggle with higher-order reading skills like blending and reading for understanding.

FIGURE 3.4: Prevalence of Zero Scores across Four Reading Assessment Tasks at Endline



Legend of statistical significance of differences between project and comparison schools: *** < 0.001, ** < 0.01, * < 0.05

3.3 Comparisons by Gender

Room to Read was also interested in understanding the role gender plays alongside the program intervention. Table 3.2 below provides the assessment results disaggregated by gender for project and comparison school children. **Data show that the Literacy Program benefitted both boys and girls.** Across all tasks, both girls and boys from project schools scored higher and experienced greater gains than their counterparts in comparison schools ($p < 0.001$).

Though both boys and girls benefited from the Literacy Program, **the benefits for girls versus boys varied by assessment task**. Project school girls made significantly greater gains than project school boys on the non-word reading task ($p < 0.05$) and the oral reading fluency task ($p < 0.001$). On the letter sounding and comprehension tasks, the differences in gains between project girls and project boys were not statistically significant. As shown in Figures 3.5 and 3.6, the effect sizes were large for both boys and girls, though they were slightly higher for boys than for girls on most tasks.

TABLE 3.2: Comparison of Average Scores by Gender

Assessment Task	Group	BASELINE n	BASELINE Mean	BASELINE SD	MIDLINE (END OF GRADE 1) n	MIDLINE (END OF GRADE 1) Mean	MIDLINE (END OF GRADE 1) SD	ENDLINE (END OF GRADE 2) n	ENDLINE (END OF GRADE 2) Mean	ENDLINE (END OF GRADE 2) SD	2-Year Gains	Adjusted Difference in 2-Year Gains[†]
Letter sounding fluency	Project boys	483	2.9	6.6	457	27.4	18.8	449	50.5	24.7	+47.6	**D1: +21.7*** **
(letters per minute)	Comparison boys	432	2.4	7.1	427	15.3	16.8	418	28.3	22.1	+25.9	**D2: +19.6*** **
	Project girls	499	3.1	6.6	462	29.8	19.5	484	52.4	24.8	+49.3	**D3: 1.7**
	Comparison girls	461	1.6	5.2	414	15.3	16.8	464	31.3	23.3	+29.7	
Non-word reading	Project boys	483	0.3	1.4	457	6.3	7.4	449	15.4	10.6	+15.1	**D1: +8.5*** **
(words per minute)	Comparison boys	432	0.3	1.7	427	2.6	5.0	418	7.0	8.8	+6.7	**D2: +8.9*** **
	Project girls	499	0.3	1.4	462	7.0	7.7	484	16.9	11.0	+16.6	**D3: 1.4***
	Comparison girls	461	0.2	1.3	414	2.5	4.9	464	7.8	8.9	+7.6	
Oral reading fluency	Project boys	483	0.5	3.1	457	7.7	11.0	449	29.7	23.8	+29.2	**D1: +19.0*** **
(words per minute)	Comparison boys	432	0.2	1.1	427	2.2	5.0	418	10.8	17.0	+10.6	**D2: +21.9*** **
	Project girls	499	0.3	1.4	462	9.3	12.8	484	34.6	26.1	+34.3	**D3: +4.4*** **
	Comparison girls	461	0.2	1.4	414	2.2	5.3	464	12.3	16.9	+12.1	
Reading comprehension	Project boys	483	0.0	0.2	457	1.0	1.3	449	2.3	1.7	+2.2	**D1: +1.2*** **
(questions answered correctly)	Comparison boys	432	0.0	0.1	427	0.3	0.8	418	1.1	1.5	+1.0	**D2: +1.3*** **
	Project girls	499	0.0	0.2	462	1.0	1.3	484	2.4	1.6	+2.3	**D3: +0.1**
	Comparison girls	461	0.0	0.1	414	0.3	0.8	464	1.0	1.4	+1.0	

D1: Adjusted differences in gains from baseline to endline between boys in project schools and boys in comparison schools.
D2: Adjusted differences in gains from baseline to endline between girls in project schools and girls in comparison schools.
D3: Adjusted differences in gains from baseline to endline between boys in project schools and girls in project schools.
[†] Adjusted difference in gains reports the coefficient of the regression analysis conducted with random effects at the school level and age and oral language spoke at home as covariates. Additionally, version of passage used during the tests were included in the regression analysis for reporting results for oral reading fluency and reading comprehension respectively.
Legend of statistical significance: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

FIGURE 3.5: Adjusted Effect Sizes – Project boys vs. Comparison boys

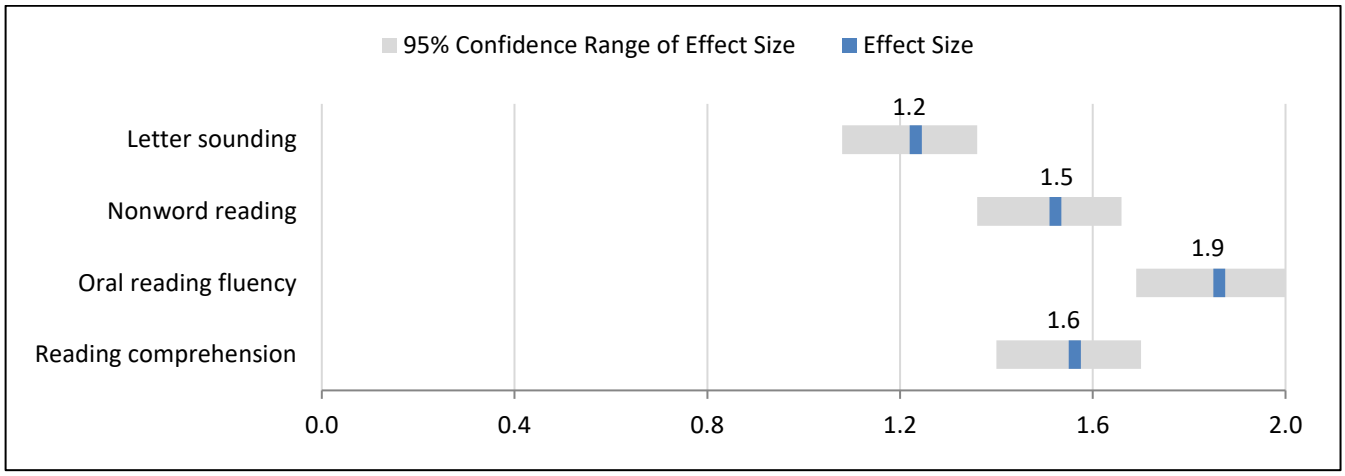
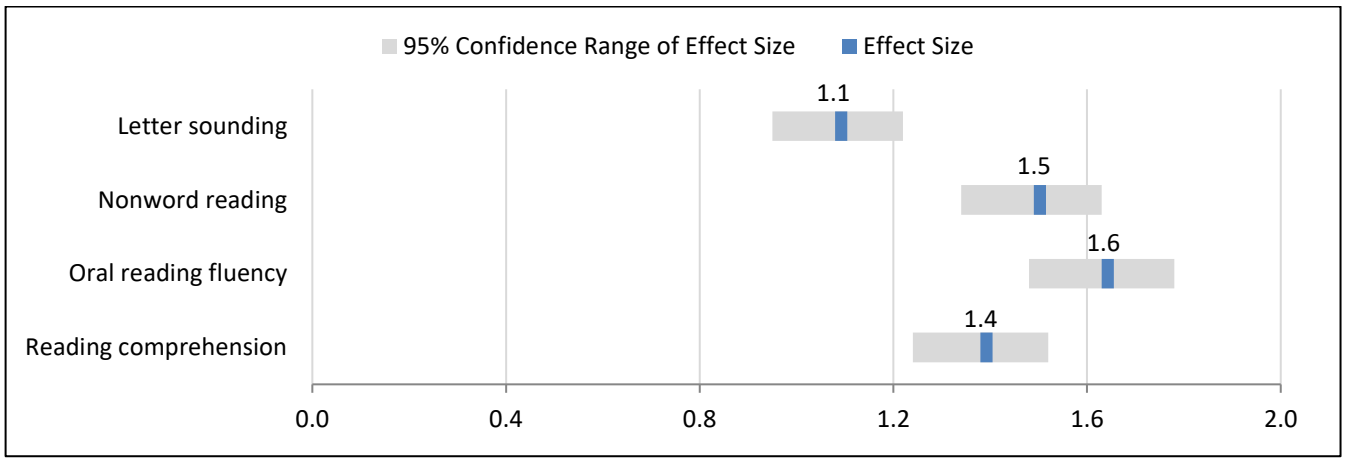


FIGURE 3.6: Adjusted Effect Sizes – Project girls vs. Comparison girls



3.4 Comparison across States

Analysis of learning assessments data across three states (Chhattisgarh, Rajasthan and Uttarakhand) revealed some interesting trends which could be useful for Room to Read's India Country and State program implementation teams.

- *Gains in reading skills after two-years of intervention in each state*: Reading assessment results for the states are presented in Annex D. In each of state, gains in reading skills for children from project schools were significantly greater than children from comparison schools across all four assessments ($p < 0.001$).In each state, maximum gains were observed for letter sounding fluency tests followed by oral reading and then non-word reading tests, respectively. Gains in reading comprehension were uniformly lowest within each of the states. Also, on each of the four assessments, biggest impact of the Room to Read's Literacy Program was observed for the children in Rajasthan.
- *Comparative analyses of reading skills at endline across states*: In Figures 3.7-3.10 below, results from state-level analyses at endline are presented. On the letter sounding, non-word reading, and oral fluency tests, children in project schools in Chhattisgarh scored significantly higher than their project school counterparts from Rajasthan and Uttarakhand ($p < 0.05$) at endline. On the reading comprehension task,

project schools children in Chhattisgarh at endline scored significantly higher than their counterparts in Rajasthan ($p < 0.001$); however, differences in the scores between Chhattisgarh and Uttarakhand on this task were not statistically significant. Notably, children in comparison schools in Chhattisgarh performed better than their counterparts in comparison schools in Rajasthan and Uttarakhand at endline. On all four reading assessment tasks, differences in scores between Chhattisgarh and Rajasthan or Uttarakhand were statistically significant for comparison school children ($p < 0.001$) at endline.

FIGURE 3.7: Letter Sounding Fluency Results across States at Endline

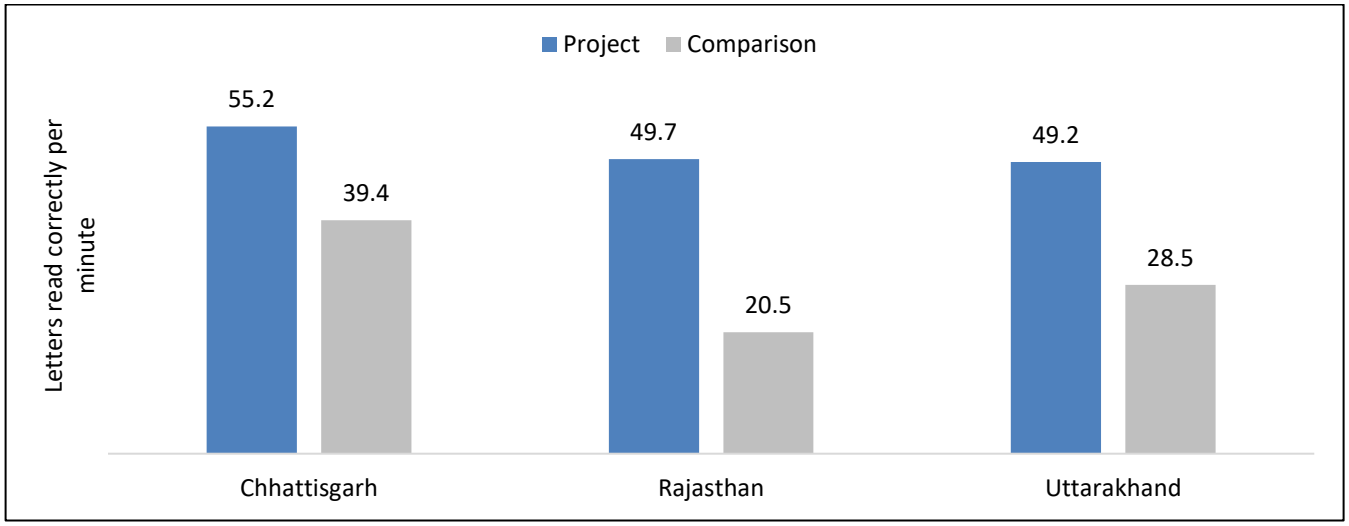


FIGURE 3.8: Non-Word Reading Results across States at Endline

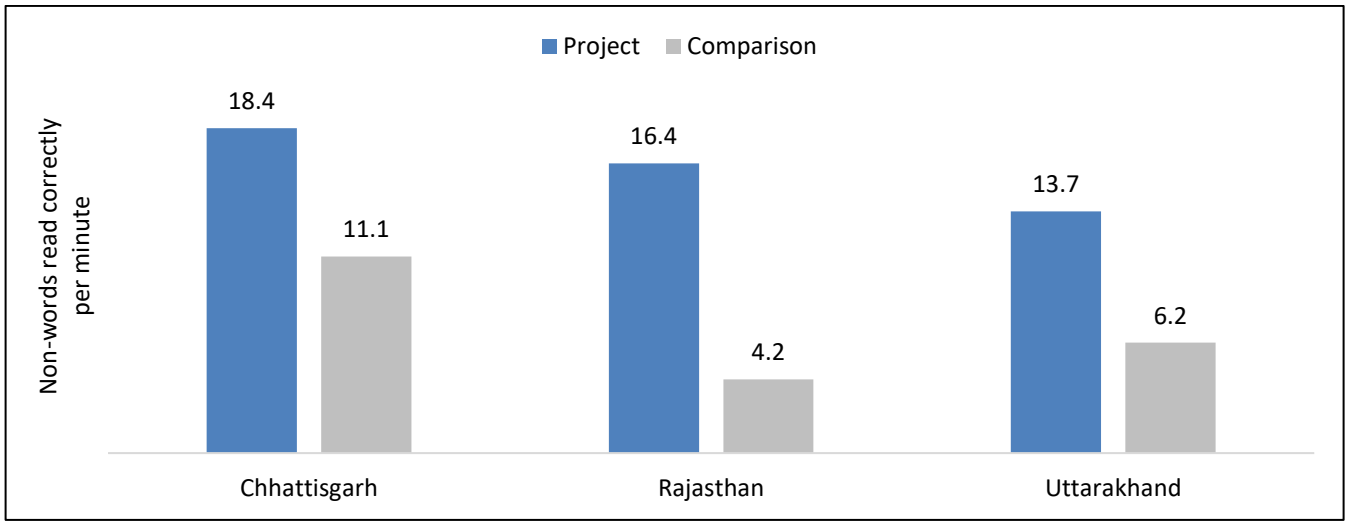


FIGURE 3.9: Oral Reading Fluency Results across States at Endline

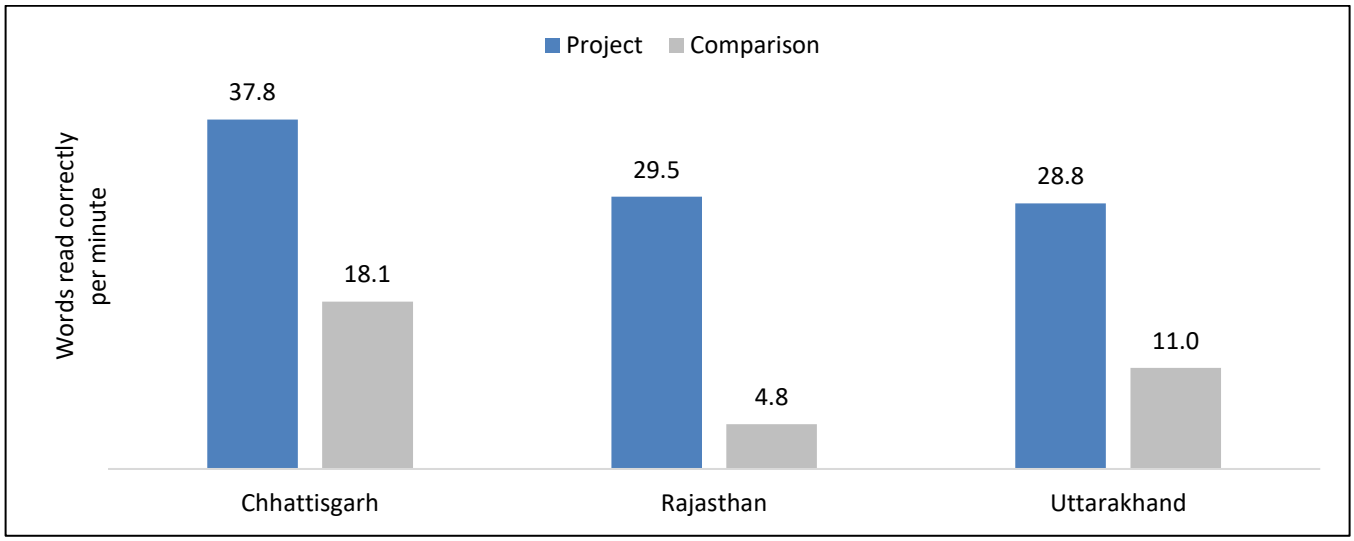
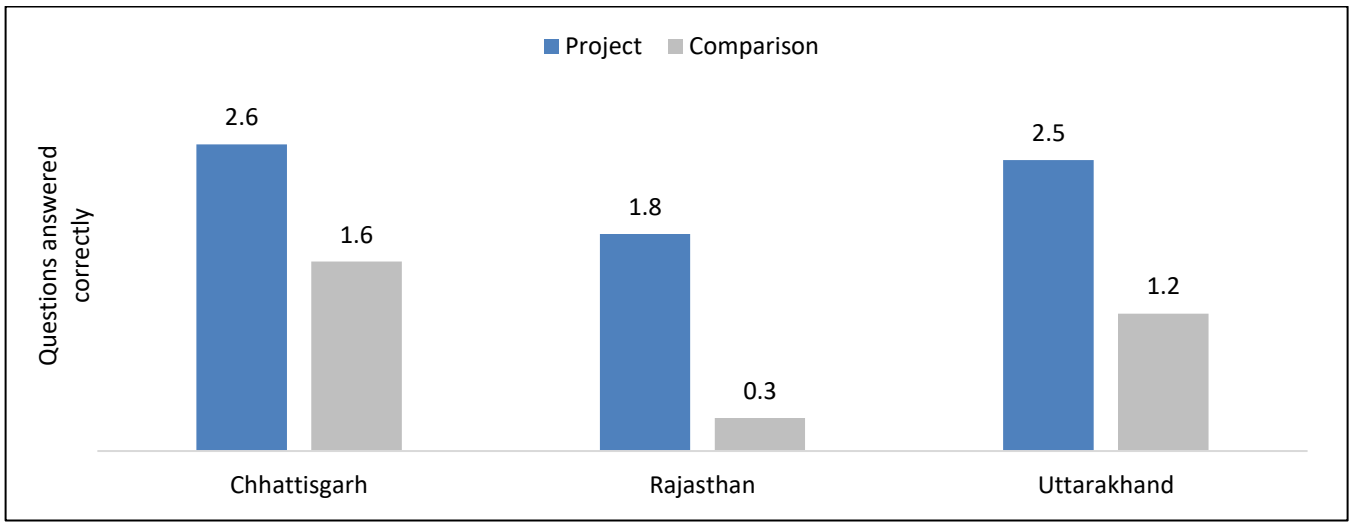


FIGURE 3.10: Reading Comprehension Results across States at Endline



In addition to the average scores across assessment tasks, we also examined state-by-state results for various benchmarks at endline (See Figure 3.11). **Consistent with the state-by-state comparison of mean scores reported in Figures 3.7-3.10, children in Chhattisgarh performed better than the other two states at endline on these benchmarks.** The percentages of project school children reading at least 20 words per minute (emergent readers) and 45 words per minute (fluent readers) were higher in Chhattisgarh than in Rajasthan or Uttarakhand ($p < 0.001$) at endline. As shown in Figure 3.12, the percentages of project school children answering at least 60 percent and 80 percent of the comprehension questions correctly were highest in Chhattisgarh and lowest in Rajasthan, respectively.

FIGURE 3.11: Emergent Reader and Fluent Reader Results for Project School Children at Endline across States

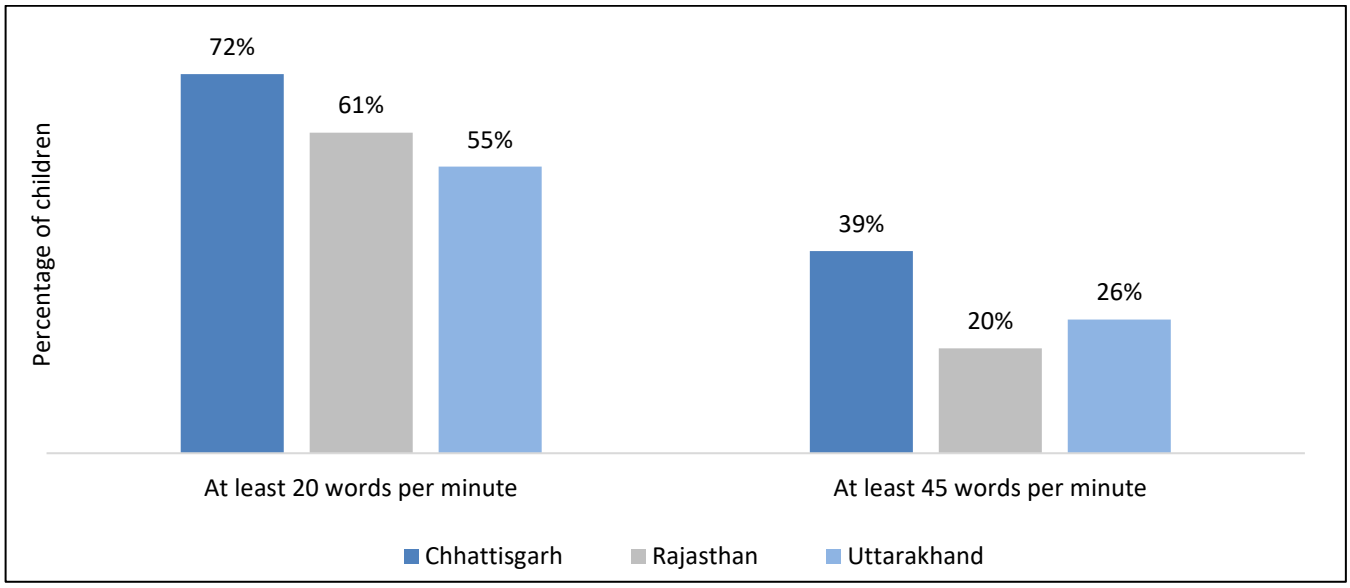
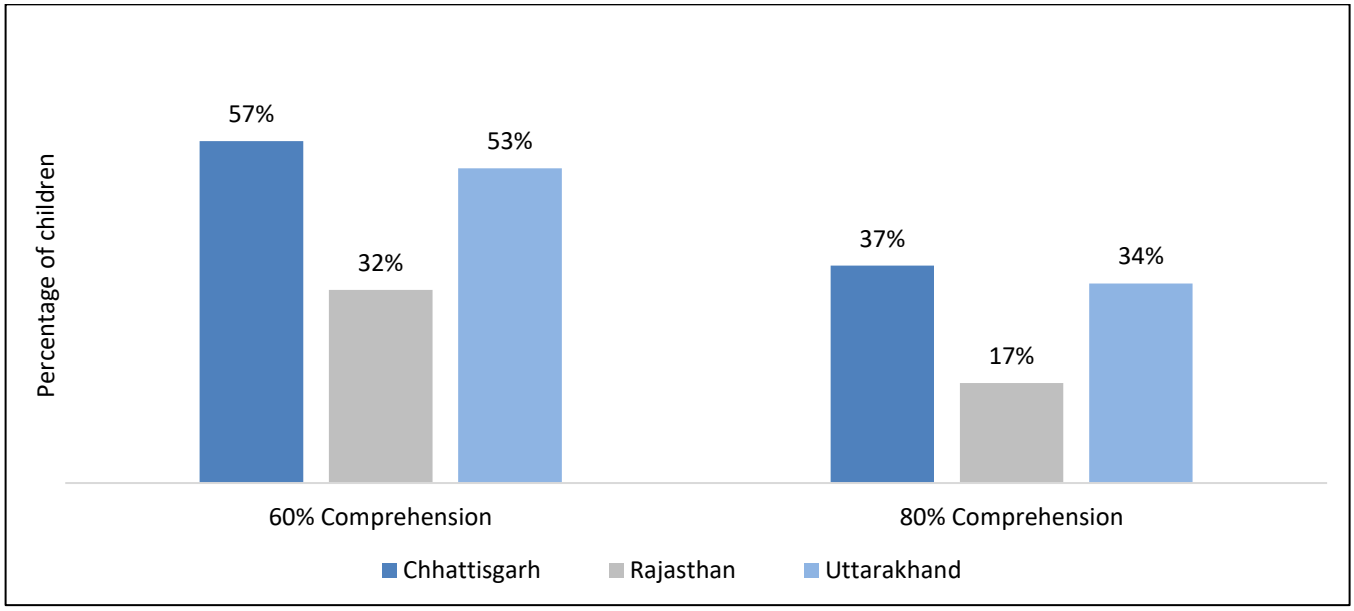


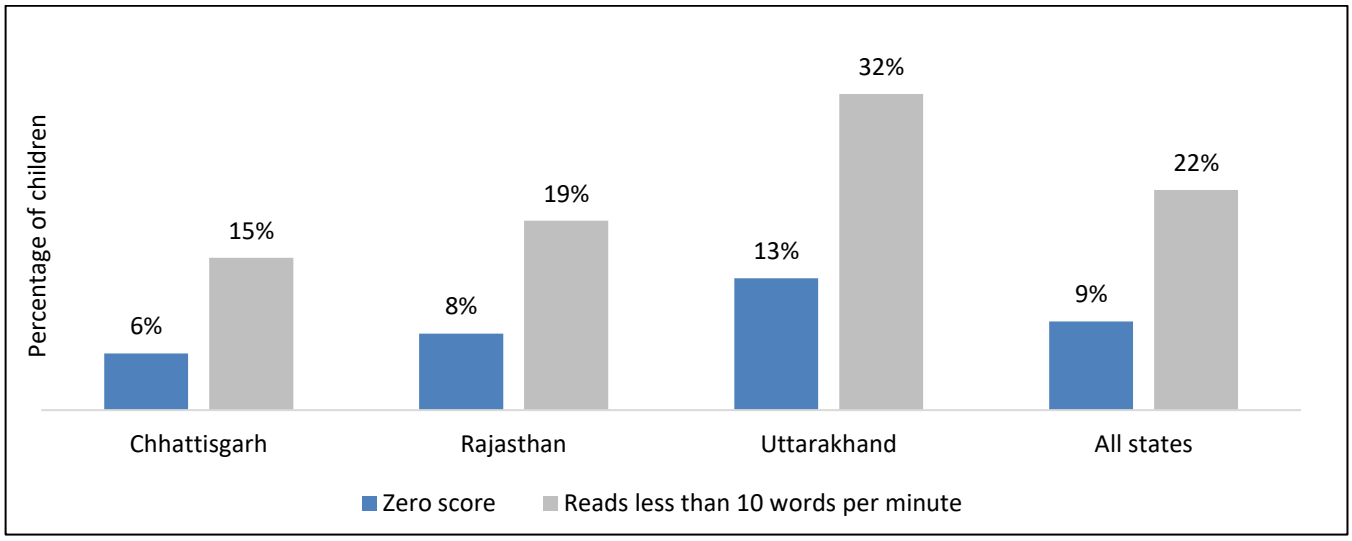
FIGURE 3.12: Percentage of Project School Children Achieving 60 and 80 percent Comprehension at Endline across States



3.5 Characteristics of Struggling Readers

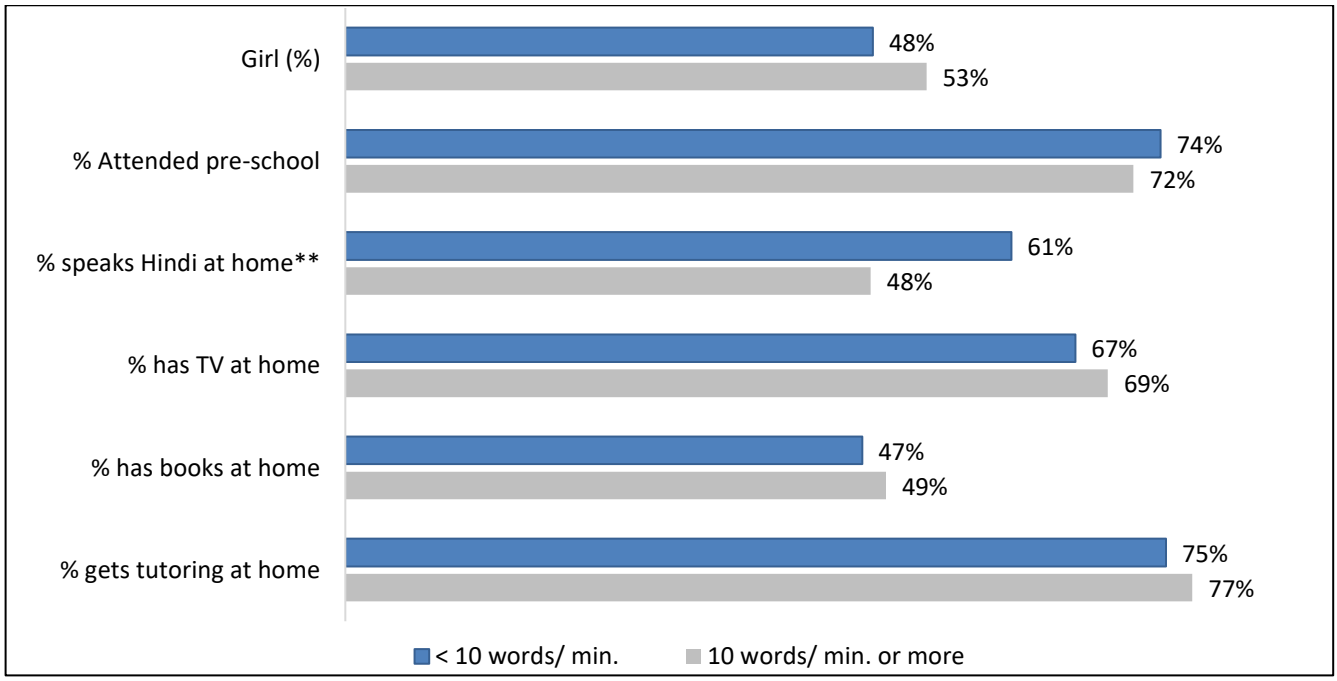
Although the evaluation showed that most project school children made significant improvements after two years of the Literacy Program intervention, nine percent of the project school children were still unable read a single word (See Figure 3.13). Distribution of children who were unable to read a single word was statistically different across states ($p < 0.01$) Moreover, 22 percent of the project school children could read less than 10 words per minute. Again, we have found statistically significant difference across states in the distribution of children who could not even read 10 words per minute ($p < 0.001$). Uttarakhand had the highest percentage of struggling readers amongst these three states as per each of these indicators mentioned here.

FIGURE 3.13: Percentage of Project School Children with Low Oral Reading Fluency at Endline across States



It is important for Room to Read to know more about these struggling readers so that we can improve program effectiveness. In Figures 3.14, we present the background characteristics of struggling readers (children who could read less than 10 words per minute) versus children with some reading ability (children who could read 10 words or more per minute) within project schools. The only child background characteristic on which children from the two groups differed ($p < 0.01$) was oral language spoken at home, as the percentage of children who spoke Hindi at home was higher among those who read less than 10 words per minute. Moreover, we have found that the differences in average oral fluency scores between Hindi and non-Hindi speaking children in project schools was not statistically significant. These findings run contrary to the assumption that Hindi-speaking children should demonstrate better reading skills than non-Hindi-speaking children on a Hindi reading assessment.

FIGURE 3.14: Characteristics of Children with Low Oral Reading Fluency in Project Schools at Endline



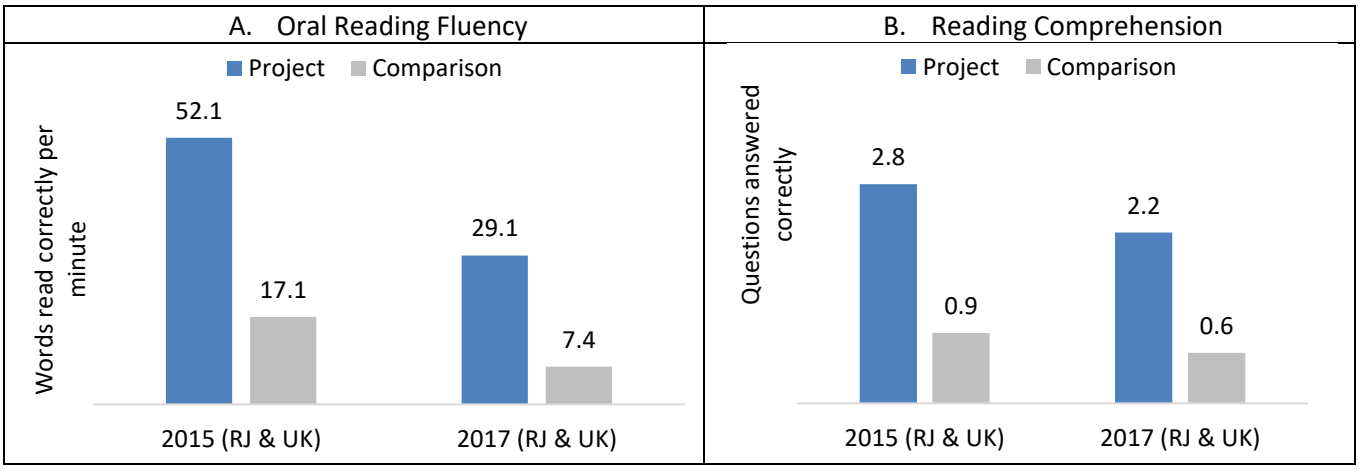
Legend of statistical significance of differences between children with low (< 10 words/ min.) and children with some reading ability (10 words/ min or more.): *** < 0.001, ** < 0.01, * < 0.05

4 Context and Limitations

4.1 Context

Previous reading skills evaluation results in India also provide context with which to interpret the 2017 endline evaluation results. Figures 4.1 below compares the 2017 end-of-Grade 2 Hindi Literacy Program evaluation results for oral reading fluency and comprehension tests with results of a similar evaluation conducted to measure the impact of the earlier version of the Literacy Program intervention in 2015. The comparison shows overall project schools results from 2015 were better than the overall project school results from 2017 for the two states (Rajasthan and Uttarakhand) that were common in the two evaluations. However, the 2017 eveluation results for children in project schools were better relative to comparison schools. Interpretation of this comparison is limited, as the programs assessed were different in design[5] and implementation, the assessments were conducted several years apart in different locations[6], and the assessment tasks may not have been of the same difficulty level.

FIGURE 4.1: End-of-Grade 2 Oral Reading Fluency and Comprehensions Results in Hindi Language Program by Evaluations



4.2 Limitations

Interpretation of the above results should include the limitations of the evaluation design. One such limitation was comparability. The validity of impact evaluation results rests on the strength of the assumption that the comparison schools, on average, are comparable to the groups of project schools amongst all observable and unobservable characteristics that may affect the outcome being evaluated. Because project schools were not randomly assigned the intervention, it may be possible that they differed from comparison schools in ways that we did not assess. As explained in *Appendix B: Research Design*, every effort was made during sampling to ensure that a comparable set of schools was chosen, and school and child-level characteristics were later analyzed to check whether any significant differences exist between project and comparison groups. As discussed in section *3.1: School and Child Background Characteristics*, there were no statistically significant difference between project and comparison groups with respect to any of the school level indicators. The only student-level background indicator for which statistically significant differences between project and comparison school children were found

[5] The 2015 evaluation was conducted for the Room to Read's erstwhile classroom-based Reading and Writing Instruction (RWI) intervention. The 2017 eveluation was done for comprehensive Literacy Program which was focused on, in addition to reading and writing instruction, improving access to reading materials through the establishment of school libraries.

[6] The 2015 eveluation was conducted in Ajmer and Tonk districts of Rajasthan and Haridwar district of Uttarakhand, respectively. In 2017, the eveluation was conducted in Jodhpur and Sirohi districts of Rajasthan and Dehradun district of Uttarakhand.

was language spoken at home among students assessed at endline. This along with age and gender of children were included in the multiple regression models as covariates to minimize the bias that can originate due to different distribution of characteristics in project and comparison schools.

5 Conclusion

Findings from the endline data indicate that the Literacy Program is having a large positive impact on reading skills. Children benefiting from the program scored higher than children from comparison schools across all reading assessment tasks. Moreover, children in program schools experienced gains in learning that were two to three times larger than those experienced by comparison school children from the beginning of Grade 1 to the end of Grade 2. The program had a notable impact on oral reading fluency, as project school children made three times as much progress as comparison school children after two years of the program.

Both boys and girls benefited significantly from the Literacy Program. Gains in learning for both in project schools were two to three times more than their counterparts in comparison schools from baseline to the end of Grade 2. However, benefits for girls versus boys varied by assessment task.

The program was effective across all three states (Chhattisgarh, Rajasthan and Uttarakhand). However, project school children in Chhattisgarh performed the best, while project school children in Rajasthan experienced the largest two-year gains relative to comparison school children.

Despite these results, room for improvement exists. For example, though the relative impact of the program was large, the oral reading fluency levels of project school children (32 words per minute, on average) was below our goal of at least 45 words per minute by the end of Grade 2. Moreover, 9 percent of project school children (and 43 percent of comparison school children) were effectively non-readers, while 21 percent of project school children (and 58 percent of comparison school children) were unable to answer even one comprehension question correctly. Uttarakhand had the highest percentage share of struggling readers across three states. These results suggest that more can be done to improve the quality and effectiveness of program delivery.

The overall program results from 2017 were not as high as the overall program results from an End-of-Grade 2 evaluation that took place in 2015. However, the 2017 program results were better relative to comparison schools.

6 Next Steps

In the coming years, Room to Read is confident that child performance will increase as teachers become more familiar with our method of teaching reading to children. Room to Read maintains that it's pacing and sequencing of phonics content, provision of high quality learning and teaching materials, and ongoing support to teachers will promote and encourage the development of student reading skills and habits. In the meantime, the Room to Read India team continues to make program adjustments that aim to improve student reading and writing outcomes in the target schools, which we believe will strengthen children's learning going forward. These key program refinements are as follows:

- Include more oral language activities in the foundation phase of the program.

- Provide more opportunities for teachers to focus on continuous remediation and revision. Specifically, teachers will have space to conduct remedial activities on the sixth day of every week for children who are struggling.
- Provide more opportunities for children to practice comprehension-based texts. This includes making independent reading time part of the regular instruction routine. During the independent reading time, we will create and encourage teachers to use story cards (i.e., simple texts with colorful illustrations) to strengthen the reading experiences of children.
- Provide greater emphasis on analysis and use of formative assessment data in Literacy Coach and teacher trainings.

Through these strategies, Room to Read hopes to achieve its goal of having all children in the program reading fluently by the end of Grade 2.

7 References

Abadzi, H. (2011). Reading Fluency Measurements in EFA FTI Partner Countries: Outcomes and Improvement Prospects. Working Paper.

Harris, T.L. & Hodges, R.E., (Eds.). (1995): The literacy dictionary: The vocabulary of reading and writing. Newark, DE: International Reading Association.

Kuhn, M., Schwanenflugel, P. J., & Meisinger, E. B. (2010). Aligning theory and assessment of reading fluency: Automaticity, prosody, and the definitions of fluency. *Reading Research Quarterly*, 45, 230-251.

Snow, C.E. (2002). Reading for understanding: Toward and R&D program for reading comprehension. RAND Corp. Available at: http://www.rand.org/pubs/monograph_reports/MR1465.html.

StataCorp. (2013). Stata Statistical Software: Release12. College Station, TX:: StataCorp LP.

Appendix A: Literacy Program Overview

The instruction component of the Literacy Program is a classroom intervention designed to complement and increase the effectiveness of the government language curriculum. The process of developing the intervention includes the completion of a scope and sequence of instruction, detailed lesson plans, classroom materials, and comprehensive teacher professional development. Literacy coaches, provide classroom support to teachers throughout the intervention.

During the research and development stage of the instruction component for any language, our country teams analyze the language curriculum and classroom instruction to determine whether all five core elements necessary in a comprehensive language curriculum are included. These elements, which are best addressed through a combination of listening, speaking, reading, and writing activities and lessons, include:

- **Phonological awareness**: Phonological awareness is knowing the sound structure of spoken language.
- **Phonics**: use of the code (sound-symbol relationships) to recognize words.
- **Vocabulary**: The knowledge of the meaning and pronunciation of words.
- **Fluency**: Fluency is determined by how quickly, accurately, and expressively someone reads, which, taken together, facilitate the reader's construction of meaning. It is demonstrated during oral reading through ease of word recognition, appropriate pacing, phrasing, and intonation. It is a factor in both oral and silent reading that can limit or support comprehension (Kuhn et al., 2010).
- **Comprehension**: A definition of reading comprehension that captures the purpose of reading is "intentional thinking during which meaning is constructed through interactions between text and reader" (Harris & Hodges, 1995, p. 207). Reading comprehension consists of three elements: the reader, the text, and the activity of reading (Snow, 2002). Writing skills are incorporated into the instructional approach through all components. In addition, teachers teach children how to write and child workbooks provide daily opportunities to practice the writing skills taught.

Appendix B: Research Design

Methodology

The impact evaluation employs a quasi-experimental design that includes children from schools that benefit from the Literacy Program (project schools) and children from schools that do not benefit from the Literacy Program (comparison schools). Data collection occurs at three points in time and follows the same cohort of children over two academic years.[7] At the beginning of Year 1, a baseline assessment was conducted with project and comparison school children who were entering Grade 1 to assess children' reading level at the beginning of program exposure. Subsequent rounds of data collection assess progress of children from the project school cohort after one and two years of the program in comparison with children from the comparison school cohort.

The structure and design of the evaluation allows for an examination of the effects of the Literacy Program on child achievement over time. The assessment results also help us understand children's reading skill strengths and weaknesses and provide Room to Read staff, classroom teachers, and administrators with information on program efficacy. These data guide program improvement strategies to ensure that children achieve the learning goals.

Sampling

The aim of sampling was to ensure that project and comparison schools were as similar as possible before the introduction of the Literacy Program. First, we assembled a list of the schools across Chhattisgarh, Rajasthan and Uttarakhand in which the Literacy Program began in 2015. We also assembled a list of non-program schools from the same blocks (or nearby blocks) as the Literacy Program schools. Next, we excluded all schools (both project and comparison) that had:

- Overall school enrollment of less than 70
- Grade 1 or Grade 2 enrollments of less than 15
- Less than three teachers in the school
- Pupil-teach ratios of more than 50:1

We randomly selected 25 project schools from each state. Since the schools were distributed across multiple blocks in each state, we stratified the sample so that the proportion of project schools selected by block matched the overall proportion of project schools by block within the state. Next, we determined the median pupil-teacher ratio of the selected project schools for each state and categorized the potential comparison schools for that state as either "small" (below the median) or "large" (above the median). We then randomly selected 25 comparison schools from each state. Wherever possible, we selected (i) the same number of comparison schools as project schools within each block and (ii) an equal number of small or large schools within each block.

During each data collection period, 15 children from Grade 2 from each school were randomly selected if they satisfied the following conditions:

- Did not have physical, sensory and significant cognitive disabilities[8]; and

[7] During each data collection point (i.e., baseline, midline), a new sample of children was randomly selected from the same cohort of children in project and comparison schools. Though the same children may be selected at multiple data collection points, Room to Read is not intentionally following the same children over the two years of the study.

[8] We were not able to identify or exclude children with learning and/or reading disabilities as such disabilities are difficult to detect in Grade 1 and 2.

- Were present on the day(s) of data collection.

In cases where the number of children present on the day of testing was less than 15, we assessed all children present. For the baseline, a total of 1,875 children were tested (982 from projects schools and 893 from comparison schools). For the midline, a total of 1,760 children were tested (919 from projects schools and 841 from comparison schools). For the endline, a total of 1,815 children were tested (933 from projects schools and 882 from comparison schools).

Reading Assessments

In this evaluation, Room to Read is assessing children' literacy skills using a version of the Early Grade Reading Assessment (EGRA)[9] that was adapted from English to India language by local experts. Room to Read used a version of the EGRA that was designed according to the expected reading levels of Grade 2. The EGRA was comprised of four common tasks:

- *Letter sounding fluency:* Ability to read letters of the alphabet without hesitation and naturally. This is a timed test that assesses automaticity and fluency of letter recognition. Children are given one minute to read 100 letters.
- *Non-word fluency:* Ability to read words that do not exist, but whose letter combinations follow the rules of the language. This task assesses the child's ability to "decode" words fluently as distinct from their ability to recognize words they have seen before. Children are given one minute to read 50 words.
- *Oral reading fluency:* Ability to read a 60-word passage that tells a story. Children are given three minutes to read the passage.
- *Reading comprehension:* Ability to answer five questions based on the passage.

Assessments were administered individually with children by external data collectors. Room to Read hired Sigma Research and Consulting Private Limited, an India-based research firm, to manage the data collection process

Assessor Training

The external research agency, Sigma, with support from Room to Read's Research, Monitoring, and Evaluation (RM&E) team and Instruction Design and Technical Support (IDTS) team, conducted the training. The four-day training included three days of in-house training which focused on sub-test structure, administration processes, and demonstration and a one-day exposure visit to a non-sample school for practice with children. At the end of the assessor training, an inter-reliability testing was conducted to select the best 20 assessors based, and 10 of them were assigned the role of supervisors.

Data Collection

Endline data collection took place in three phases: first in Chhattisgarh and then in Uttarakhand. In Chhattisgarh, data collection took place during the last week of February and first week of March 2017 at a total of 50 schools – 25 from the project group and 25 from the comparison group. In Uttarakhand, data collection took place during the second and third weeks of March 2017 in the same number of schools (50 in total – 25 project and 25 comparison). In Rajasthan, data collection took place during the second and third weeks of April 2017 in the same number of schools (50 in total – 25 project and 25 comparison). In all states, data collection was conducted by 10

[9] The EGRA was developed by RTI International in 2006. For more information, please see: https://www.eddataglobal.org/reading/index.cfm?fuseaction=pubDetail&id=929.

data collection teams, and each team was comprised of one assessor and one supervisor. The supervisors were trained and assigned additional tasks of field level management, ensuring the compliance with assessment administration process, and maintaining data quality at field. The supervisors ensured that the assessments were administered correctly and that all the necessary data were captured. A member of the Research, Monitoring, and Evaluation team also acted as a quality supervisor for the entire data collection and was in touch with each team supervisor to conduct quick reflection every day after data collection.

Data Entry

Data entry was done by a data entry operator hired by the external research agency using a data entry template developed by Room to Read. The point person for data entry was trained by RM&E staff on how to enter data from hard copies in to the Excel format. Post-data entry, 10 percent of the entered cases were randomly verified with the actual hard copies. Entered data were exported to Stata for analysis.

Data Analysis

We started the analysis by comparing the background variables of schools and children between the project and comparison groups. For the school background variables, we examined differences in mean Grade 2 enrollment, average number of in-position teachers, average language arts instruction time and average children's attendance on the day of the endline assessment using T-tests. We examined differences in type of class (mono-grade vs. multi-grade), presence of dedicated language-arts teacher and school location (urban vs. rural) using Chi-square tests. For the child background variables, we examined differences in age, gender , language spoke at home, whether the child attended pre-school, whether the child had a television and books at home, and whether the child receives tutoring at home by conducting regression analysis (linear regression for age and logistic regression for other categorical variables) with random effects at the school level. The equations included the child background variable as the dependent variable and school type (project or comparison) as the independent variable. The results of these analyses appear in *Appendix C*.

The primary aim of the data analysis was to determine if children in the project group made greater gains from baseline to endline than children in the comparison group. The analysis strategy was to compare reading levels in the two assessment periods (baseline versus endline) among the two experimental groups (project versus comparison). An impact of the program is evident if there is a greater gain from baseline to endline among the project group compared to the comparison group. This is demonstrated by a statistically significant interaction between experimental group and assessment period. To determine this, we conducted linear regression analysis with random effects at the school level and dummy variables for the assessment period, experimental group, and the interaction between the two with scores in different reading tasks (from both baseline and endline) as the dependent variable. Each regression model also included age, sex, and language spoke at home as the covariates[10]. We followed a similar procedure to analyze differences in gains by gender. We conducted separate analysis to examine differences in gains across each of the following: project school boys versus comparison school boys, project school girls versus comparison school girls, and project school boys versus project school girls. Each analysis included one of the assessment scores as the dependent variable and age and language spoke at home as the covariates.

[10] Age and sex were included in the regression model because of their known effects on children' reading performance. Language spoke at home was included because of a significant difference between project and comparison school children for this variable at endline (see *Appendix C*).

Further, we analyzed zero scores to determine the impact of the intervention on prevalence of non-readers. The analysis of zero scores is particularly appropriate when the distribution of scores is skewed towards zero (i.e., is not in a bell-shaped curve). We conducted logistic regression analysis with random effects at the school level to determine if significant differences existed between the percentage of project school vs. comparison school children registering zero scores (vs. non-zero scores). The equations included the presence or absence of a zero score as the dependent variable, type of school as the predictor, and language spoke at home, age and sex as covariates.

We also computed effect size for different outcome measures to demonstrate the magnitude of effect of intervention. Effect size is mean difference in gains between the project and comparison schools divided by the pooled standard deviation of scores in project and comparison schools. An effect size is exactly equivalent to a "Z-score" of a standard normal distribution. For example, an effect size of "+1" means that the score of the average child in the project school is "+1" standard deviations above the average person in the comparison school, and hence exceeds the scores of 84 percent of the comparison group.

We examined the differences across states in (a) average reading scores on all four tasks, (b) average shares of emergent readers and fluent readers, and (c) average shares of various types of struggling readers (children who could not read a single word and children who could read up to 10 words per minute) through Analysis of Variance (ANOVA) tests.

Appendix C: School and Child Background Characteristics

TABLE C.1: Background Characteristics of Sample Schools

	Project		Control	
	n	% or Mean (SD)	n	% or Mean (SD)
Schools	75	-	75	-
States				
Chhattisgarh	25	33%	25	33%
Rajasthan	25	33%	25	33%
Uttarakhand	25	33%	25	33%
District				
Dehradun (Uttarakhand)	25	33%	25	33%
Jodhpur (Rajasthan)	17	23%	17	23%
Raipur (Chhattisgarh)	25	33%	25	33%
Sirohi (Rajasthan)	8	11%	8	11%
Location				
Rural	10	13%	11	15%
Urban	65	87%	64	85%
Grade 2 Enrolment	75	24.0 (14.1)	75	23.9 (14.0)
Total teachers in the school	75	4.9 (2.2)	75	4.4 (1.9)
Percentage of children present on day of Endline assessment	75	72%	75	71%
Designated minutes of reading instruction per day	75	48.8 (16.7)	75	45.0 (11.5)
Classroom type				
Mono-grade	40	53%	48	64%
Multi-grade	35	47%	27	36%
Designated language arts teacher present in the school	68	91%	62	83%

Legend of statistical significance: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

TABLE C.2: Background Characteristics of Sample Children

	Project		Control	
	n	% or Mean (SD)	n	% or Mean (SD)
Age				
Baseline	982	5.9 (1.0)	893	5.9 (0.8)
Endline	933	7.6 (1.2)	882	7.7 (1.1)
Gender - Baseline				
Boys	483	49%	432	48%
Girls	499	51%	461	52%
Gender - Endline				
Boys	449	48%	418	47%
Girls	484	52%	464	53%
Home language - Baseline				
Marwari	327	33%	338	38%
Chhattisgarhi	320	33%	316	35%
Hindi	335	34%	234	26%
Other	0	0%	5	1%
Home language - Endline				
Hindi***	473	51%	362	41%
Marwari	261	28%	305	35%
Chhattisgarhi	199	21%	215	24%
Attended preschool				
Baseline	645	66%	620	69%
Endline	675	72%	584	66%
TV at home (endline only)	642	69%	563	64%
Books at home (endline only)	455	49%	428	49%
Receive tutoring at home (endline only)	715	77%	618	70%

Legend of statistical significance: *** p < 0.001, ** p < 0.01, * p < 0.05.

TABLE C.3: Correlation between different Reading Tasks

	Letter sounding	Non-word reading	Oral reading fluency	Reading comprehension
Letter sounding	1.000			
Non-word reading	0.855***	1.000		
Oral reading fluency	0.808***	0.925***	1.000	
Reading comprehension	0.762***	0.775***	0.790***	1.000

Legend of statistical significance: *** p < 0.001, ** p < 0.01, * p < 0.05.

Appendix D: Reading Assessment Results by State

TABLE D.1: Reading Assessment Results – Chhattisgarh

Assessment Task	Group	BASELINE n	BASELINE Mean	BASELINE SD	MIDLINE (END OF GRADE 1) n	MIDLINE (END OF GRADE 1) Mean	MIDLINE (END OF GRADE 1) SD	ENDLINE (END OF GRADE 2) n	ENDLINE (END OF GRADE 2) Mean	ENDLINE (END OF GRADE 2) SD	2-Year Gains	Adjusted Difference in 2-Year Gains over Comparison[†]
Letter sounding fluency	Project	345	2.7	5.7	342	35.2	17.5	334	55.2	24.0	+52.5	15.7***
(letters per minute)	Comparison	351	2.7	7.2	340	23.9	16.7	344	39.4	22.3	+36.8	
Non-word reading	Project	345	0.2	1.1	342	9.3	8.1	334	18.4	10.6	+18.2	7.2***
(non-words per minute)	Comparison	351	0.2	1.2	340	4.4	6.0	344	11.1	9.6	+10.9	
Oral reading fluency	Project	345	0.2	1.2	342	10.6	11.5	334	37.8	25.4	+37.6	19.4***
(words per minute)	Comparison	351	0.2	1.2	340	4.0	6.4	344	18.1	18.0	+17.9	
Reading comprehension	Project	345	0.0	0.1	342	1.3	1.4	334	2.6	1.5	+2.6	1.0***
(questions answered correctly)	Comparison	351	0.0	0.2	340	0.6	1.0	344	1.6	1.6	+1.6	

Legend of statistical significance of differences between project and comparison schools: *** < 0.001, ** < 0.01, * < 0.05.

[†] Adjusted difference in gains reports the coefficient of the regression analysis conducted with random effects at the school level and age, sex and language spoken at home as covariates. Additionally, version of passage used during the tests were included in the regression analysis for reporting results for oral reading fluency and reading comprehension, respectively.

TABLE D.2: Reading Assessment Results – Rajasthan

Assessment Task	Group	BASELINE n	BASELINE Mean	BASELINE SD	MIDLINE (END OF GRADE 1) n	MIDLINE (END OF GRADE 1) Mean	MIDLINE (END OF GRADE 1) SD	ENDLINE (END OF GRADE 2) n	ENDLINE (END OF GRADE 2) Mean	ENDLINE (END OF GRADE 2) SD	2-Year Gains	Adjusted Difference in 2-Year Gains over Comparison[†]
Letter sounding fluency	Project	310	2.1	5.0	301	27.8	20.4	274	49.7	25.3	+47.6	29.0***
(letters per minute)	Comparison	338	1.7	5.4	200	9.7	13.7	315	20.5	20.0	+18.7	
Non-word reading	Project	310	0.2	1.3	301	5.1	7.0	274	16.4	9.9	+16.2	12.2***
(non-words per minute)	Comparison	338	0.3	2.0	200	1.1	2.7	315	4.2	6.9	+3.9	
Oral reading fluency	Project	310	0.3	1.6	301	8.0	12.4	274	29.5	22.6	+29.2	24.4***
(words per minute)	Comparison	338	0.0	0.3	200	0.9	3.6	315	4.8	11.5	+4.8	
Reading comprehension	Project	310	0.0	0.1	301	1.2	1.4	274	1.8	1.5	+1.8	1.5***
(questions answered correctly)	Comparison	338	0.0	0.1	200	0.2	0.7	315	0.3	0.8	+0.3	

Legend of statistical significance of differences between project and comparison schools: *** < 0.001, ** < 0.01, * < 0.05.

[†] Adjusted difference in gains reports the coefficient of the regression analysis conducted with random effects at the school level and age, sex and language spoken at home as covariates. Additionally, version of passage used during the tests were included in the regression analysis for reporting results for oral reading fluency and reading comprehension, respectively.

TABLE D.3: Reading Assessment Results – Uttarakhand

Assessment Task	Group	BASELINE n	BASELINE Mean	BASELINE SD	MIDLINE (END OF GRADE 1) n	MIDLINE (END OF GRADE 1) Mean	MIDLINE (END OF GRADE 1) SD	ENDLINE (END OF GRADE 2) n	ENDLINE (END OF GRADE 2) Mean	ENDLINE (END OF GRADE 2) SD	2-Year Gains	Adjusted Difference in 2-Year Gains over Comparison[†]
Letter sounding fluency	Project	327	4.2	8.5	276	21.3	17.2	325	49.2	24.6	+45.0	17.8***
(letters per minute)	Comparison	204	1.3	5.2	301	9.3	14.4	223	28.5	21.4	+27.2	
Non-word reading	Project	327	0.5	1.8	276	5.1	6.5	325	13.7	11.2	+13.1	7.1***
(non-words per minute)	Comparison	204	0.2	1.1	301	1.5	4.0	223	6.2	7.9	+6.0	
Oral reading fluency	Project	327	0.8	3.7	276	6.3	11.7	325	28.8	25.9	+28.0	17.4***
(words per minute)	Comparison	204	0.3	2.0	301	1.0	3.5	223	11.0	18.0	+10.6	
Reading comprehension	Project	327	0.1	0.3	276	0.5	1.0	325	2.5	1.7	+2.4	1.3***
(questions answered correctly)	Comparison	204	0.0	0.2	301	0.1	0.5	223	1.2	1.5	+1.1	

Legend of statistical significance of differences between project and comparison schools: *** < 0.001, ** < 0.01, * < 0.05.

[†] Adjusted difference in gains reports the coefficient of the regression analysis conducted with random effects at the school level and age, sex and language spoken at home as covariates. Additionally, version of passage used during the tests were included in the regression analysis for reporting results for oral reading fluency and reading comprehension, respectively.