World Change Starts with Educated Children.®

Impact Evaluation of the Literacy Program-Partnership Approach under Scaling up Early Reading Intervention(SERI) funded by USAID

2018 Endline Evaluation Report for Chhattisgarh and Uttarakhand

Pinaki Joddar

July 2018

In September 2015, Room to Read entered into an agreement with the United States Agency for International Development for implementing a large-scale, innovative, early-grade reading program aimed at benefiting children in the government primary schools in the states of Chhattisgarh and Uttarakhand. In addition to direct implementation in select government primary schools (defined as demonstration approach), Room to Read is also collaborating with the state governments since 2016 to assist them in rolling out an effective model based on its Literacy Program to improve reading outcomes among early grade children (defined as partnership approach). Alongside, Room to Read began a two-year study to assess the impact of this Literacy Program under partnership approach by contrasting the changes in the reading skills of children in a sample of school benefitting from the program for two years (project schools) with the changes in the same in a sample of matching comparison schools not getting benefits from the program during the same time period. Room to Read collected data on reading skills in project schools before the launch of the program from start-of-Grade 1 children in August 2016 (baseline) and after two years of program from end-of-Grade 2 children in Feb-March 2018 (endline); reading skills data from comparison school children was also collected following the same timeline. The results of the impact evaluation showed that children in project schools experienced significantly greater gains in reading skills than children in comparison schools from the beginning of Grade 1 to the end of Grade 2. Children who benefitted from the program could read more fluently than children who did not benefit from the program. In addition, the program had a positive effect on reading skills on both boys and girls. A key concern for program will be to reach those children who are still performing at low level.

This page is left blank intentionally

Table of Contents

1 Executive Summary

In September 2015, Room to Read entered into an agreement with the United States Agency for International Development (USAID) for implementing a large-scale, innovative, early grade reading program aimed at benefiting children in government primary schools in India. Known as the Scaling up Early Reading Intervention (SERI), the project aims to demonstrate an effective model towards improving reading outcomes among primary grade children and an innovative approach for scaling NGO-led interventions through the government system. The scale up is based on a scaffold approach which includes: (i) *Demonstration*: Demonstration of the model through direct implementation by Room to Read in select government primary schools in some districts; (ii) *Partnership*: Collaborate closely to assist the state government in expansion of the model across entire district; and (ii) *Scale-up*: Handing over the model to the government for replication in other districts.

The foundation for SERI is Room to Read's Literacy Program, a multi-faceted intervention that works in collaboration with local governments to strengthen child reading skills and habits in the primary grades. Room to Read's Literacy Program consists of two core components: (i) literacy instruction and (ii) school libraries and quality reading materials. The two components of the Program strongly complement each other and together promote a holistic approach to literacy development, thus ensuring that children can develop reading skills and reading habits. To support Grade 1 and 2 children and teachers, the instruction component of the Literacy Program provides: (i) child books with content that directly targets the development of key reading and writing skills, (ii) a teacher's guide containing easy to follow lesson plans and instructions for every lesson, and (iii) teacher training and ongoing teacher support. Through the library component of the Program, Room to Read establishes school libraries, fills libraries with engaging reading materials, and trains school personnel on how to best manage the library's resources and facilitate reading activities with children.

In 2016, Room to Read began a two-year impact evaluation of the Literacy Program under the partnership approach funded by SERI in Chhattisgarh and Uttarakhand that was operationalized in 1,010 government primary schools at the beginning of 2016-17 academic year. The evaluation aims to assess the impact of the Program by contrasting the changes in the reading skills of children in schools benefiting from the intervention for two years (project group) with the changes in a matching set of schools during the same time period but not benefiting from the intervention (comparison group). The evaluation included 74 project schools (38 units from Chhattisgarh and 36 units from Uttarakhand) and 72 comparison schools (38 units from Chhattisgarh and 34 units from Uttarakhand). Room to Read hired a survey agency to collect data on children's reading skills from these schools using a version of the Early Grade Reading Assessment (EGRA) that was adapted to Hindi by local experts.

In August 2016, Room to Read conducted the baseline assessments before the launch of the Program and found that children from project schools and comparison schools entered Grade 1 with similarly low reading skills. In February-March 2018, we conducted the final round of assessments (endline) in the same schools included in the baseline with different cross-section of children at the end-of-Grade 2 to measure the impact of the Program after two academic years. Data revealed that the program was having a large positive impact on the development of children's reading skills. Children benefiting from the Program scored higher than children from comparison schools across all reading assessment tasks at endline. **Most importantly, children in project schools also experienced significantly greater gains in reading skills than children in comparison schools from the beginning of Grade 1 to the end of Grade 2.** By the end of Grade 2, children from program schools could, on average, read 37 correct words per minute on a test of oral reading fluency. By contrast, children from comparison schools could

read only 18 correct 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.

In Chhattisgarh and Uttarakhand, average oral reading fluency of children in project schools after two years of intervention were 35 and 40 correct words per minute receptively (the corresponding values in the comparison schools were 14 and 24 correct words per minute respectively). **In each state, 2-year gains made by children in project schools on oral reading fluency and reading comprehension were significantly higher than their respective counterparts in comparison schools. Similarly, improvement in reading skills made by both boys and girls in project schools from baseline to endline were significantly greater than the improvements made by their respective counterparts in comparison schools.**

Overall, these results establishes that the large positive effect on reading skills observed in the project schools in Chhattisgarh and Uttarakhand was attributable to the Hindi Literacy Program under partnership approach. However, the results also point to areas for improvement. For example, the overall reading fluency levels of project school children (37 correct words per minute, on average) was below our goal of at least 45 correct words per minute by the end of Grade 2. Moreover, seven percent of project school children were effectively non-readers, while 16 percent of project school children were unable to answer even one reading comprehension question correctly. These results suggest that more can be done to improve the quality and effectiveness of program delivery.

2 Introduction

2.1 Literacy Program under Partnership Approach

In September 2015, Room to Read entered into an agreement with the United States Agency for International Development (USAID) for implementing a large-scale, innovative, early grade reading program aimed at benefiting children in government primary schools in India. Known as the Scaling up Early Reading Intervention (SERI), the project aims to demonstrate an effective model towards improving reading outcomes among primary grade children and an innovative approach for scaling NGO-led interventions through the government system. The scale up is based on a scaffold approach which includes:

(i) *Demonstration*: Demonstration of the model through direct implementation by Room to Read in select government primary schools in some districts;
(ii) *Partnership*: Collaborate closely to assist the state government in expansion of the model across entire district; and
(iii) *Scale-up*: Handing over the model to the government for replication in other districts.

The foundation for SERI is Room to Read's Literacy Program[1], a school-based intervention that seeks to develop children's reading skills and reading habits in the early primary grades. The program includes two main components: (1) *Instruction*: reading and writing instruction for children in Grades 1 and 2; and (ii) *Library and Quality Reading Materials (QRM)*: access to quality reading materials through the establishment of school libraries. For the instruction component, the Literacy Program team in India worked with the Department of Education to design a supplementary program that provides a strong foundation in reading and writing skills for all early primary grade 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 Room to Read appointed Literacy Coaches/ Facilitators (see *Appendix A* for a more detailed description of the instruction component of the Literacy Program).

As part of the partnership approach of SERI, Room to Read launched the Literacy Program at 500 schools in Baloda Bazar district in Chhattisgarh and 510 schools in Champawat district in Uttarakhand at the beginning of the 2016-17 academic year.

2.2 Evaluation of the Program

To determine the effect of the intervention on children's reading proficiency, Room to Read initiated a two-year (2016-18) impact evaluation of the Hindi Literacy Program under the partnership approach in these two states. The objectives of the impact evaluation are:

- Determine whether the Program is having an impact on children's reading skills after two academic years;
- 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

[1] Alternatively referred as "Program" in this documents.

- Identify reading skills that could be better supported by the Program and determine how to improve these reading skills quickly and effectively.

Room to Read with support from independent data collectors is conducting this impact evaluation.

2.3 Evaluation Methodology

As the Literacy Program under partnership approach is getting implemented in a set of government primary schools that were selected in a non-random manner using pre-decided criteria, traditional and preferred methods for impact evaluation (like randomized control trial) cannot be used here. Consequently, this impact evaluation employs a quasi-experimental design that includes collection of reading skills data from (i) children in a sample of schools benefiting from the Program (Project group); and (ii) children in a sample of schools not benefiting from the Program (comparison group) but having characteristics similar to the schools in the project group. It was planned that data collection would occur at two points in time and follows the same schools but different cross-section of children over two academic years, such as:[2]

- *Baseline*: Assesses reading skills of a sample of children at the start-of-Grade 1 in both project and comparison schools before the launch of the Program; and
- *Endline*: Assesses reading skills of a sample of children at the end-of-Grade 2 in the same project and comparison schools after two full years of implementation of the Program.

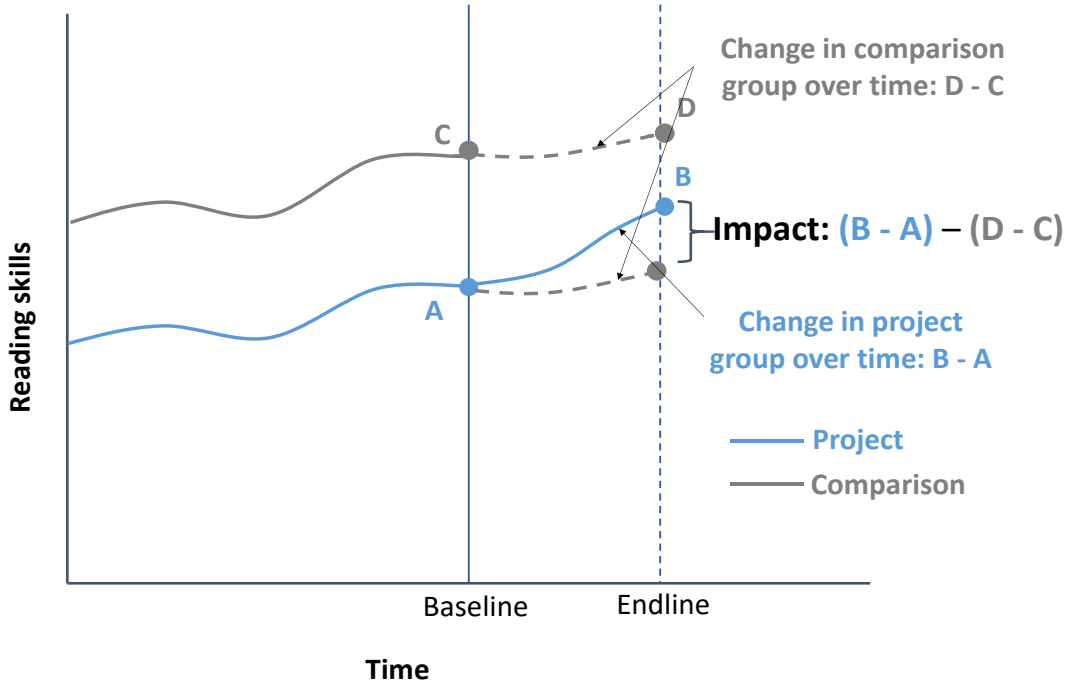
Using the these two waves of data on reading skills, the quasi-experimental methodology allows Room to Read to estimate the impact of the Program by comparing the difference between the following two categories after controlling for various school- and child-specific background characteristics that may influence their reading skills through regression analysis:

- *Category 1 (first difference)*: Changes in reading skills of children from baseline to endline in project schools; and
- Category 2 (*second difference*): Changes in reading skills of children from baseline to endline in comparison schools.

Using this specific impact evaluation technique, Room to Read will be able to, (i) estimate the extent of impact (i.e. effect size) of the Program on reading skills of children in project schools, and (ii) establish a direct causal relationship between the estimated impact on reading skills and the Program. In impact evaluation literature, this methodology is defined as “Difference-in-difference (DID)” or “Double difference (DD)” and a simple graphical presentation of the same is provided below. The evaluation results also help us to understand children’s reading skill strengths and weaknesses and provide Room to Read, classroom teachers, school administrators, state and local government, and USAID with information on Program effect.

[2] During each data collection point (i.e. baseline and endline), a new sample of children would be randomly selected in project and comparison schools. Though same children may get selected at multiple data collection points, Room to Read is not intentionally tracking the same children over the two years of this study.

Figure 2.1: Graphical presentation of the Difference-in-Difference Methodology used in the Impact Evaluation



Following the evaluation methodology described above, Room to Read initiated baseline reading skills data collection in August 2016 from the start-of-Grade 1 children in 74 project schools (38 from Chhattisgarh and 36 from Uttarakhand) and 72 comparison schools (38 from Chhattisgarh and 34 from Uttarakhand) respectively. Baseline results showed that children from program schools and comparison schools entered Grade 1 with similarly low reading skills. At the beginning of Grade 1, children from program schools could read an average of 11 letters per minute, while children from comparison schools could read an average of eight letters per minute. On a test of oral reading fluency, children in both program schools and comparison schools could correctly read an average of less than one word per minute.

As part of the final round of data collection, in February-March 2018, endline reading skills data was collected from the same schools from end-of-Grade 2 children to determine the impact of the program after two academic years. Findings of this impact evaluation of the Literacy program under partnership approach using these waves of data are presented in section *3. Results*.

(See *Appendix B* for a full description of the data analysis process.)

2.4 Sampling of Schools and Children for Evaluation

Through power calculations, it was determined that a sample size of 76 schools in Chhattisgarh (38 units each from the project and comparison group) and 70 schools in Uttarakhand (36 units from the project and 34 units from comparison group) with 10 children per school would be sufficient to detect a minimum effect size of 0.25 from the evaluation (assuming an α of 0.05, a power of 0.9, a baseline-endline correlation of 0.50, and an intra-class correlation of 0.05). Consequently, the aim of the sampling was to select required number of project and comparison schools in each state that were as similar as possible before the introduction of the Program. The following steps were taken to select this sample.

First, we assembled a list of the 1,010 project schools across Chhattisgarh and Uttarakhand in which the partnership model was launched with funding support from SERI in 2016. We also assembled a list of non-SERI potential comparisons schools from the same states as the SERI schools. In Chhattisgarh, we were able to assemble the list of non-SERI potential comparison schools from the same project district (Baloda Bazar). In Uttarakhand, sufficient number of non-SERI potential comparison schools was not available within the same project district (Champawat), so the list of potential comparison schools came from an adjoining district (Nainital).

In Chhattisgarh, we excluded all schools (both project and comparison) that had Grade 1 enrollments of less than 15 children. We then stratified the remaining project and comparisons schools by school ranking (as assigned by the state government), block, school enrollment, and pupil-teacher ratio. Next, 38 project schools were proportionately selected according to the share across stratum (within a strata required number of project schools was selected randomly). Lastly, 38 comparison school across stratum were selected so that same number of project and comparison schools were included in the sample from each strata (within a strata required number of comparison schools was selected randomly).

In Uttarakhand, school enrollments are lower, and we excluded all schools (bother project and comparison) that had Grade 1 enrollments of less than four children. Rest of the sampling process for selection of 70 schools (36 units from the project and 34 units from comparison group) in this state was same as in the case for Chhattisgarh.

During each round of data collection, we tried to randomly select and assess 10 children from each school. In cases where the number of children present on the day of assessment was less than 10, we tested all children present in the class. Children with any visibly noticeable physical, sensory and significant cognitive disabilities were excluded from the assessment.[3] During the 2016 baseline, a total of 1,116 start-of-Grade 1 children (607 in projects schools and 509 in comparison schools) were tested before that launch of the Program, out of which 755 were from Chhattisgarh and 361 were from Uttarakhand respectively. During the 2018 endline, 941 end-of-Grade 2 children (481 in projects schools and 460 in comparison schools) were tested after two years of implementation of the Program, out of which 591 were from Chhattisgarh and 350 were from Uttarakhand.

2.5 Tool used for Reading Skills Assessment

In this impact evaluation, Room to Read assessed children's reading skills using a version of the Early Grade Reading Assessment (EGRA)[4] that was adapted from English to Hindi 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 five 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 were given one minute to read 100 letters.
- *Non-word reading 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

[3] Data collectors were not able to identify or exclude children with learning and/or reading disabilities as such disabilities are difficult to detect in early grades.

[4] The EGRA was developed by RTI International in 2006. For more information, please see: https://shared.rti.org/content/early-grade-reading-assessment-egra-toolkit-second-edition.

their ability to recognize words they have seen before. Children were given one minute to read 50 non-words.

- *Oral reading fluency:* Ability to read a 60-word passage that tells a story. Children were given up to three minutes to read the passage. The evaluation used four different versions of Hindi reading passages in the actual tests, which were randomly used across children.
- *Reading comprehension:* Ability to answer five comprehension questions based on the passage.

Assessments were administered individually with children by an external data collection agency that was recruited and trained by Room to Read. Room to Read hired Sigma Research and Consulting Private Limited, an India-based survey firm, to manage the data collection.

(See *Appendix C* for a full description of the training of assessors and data collection.)

3 Results

3.1 School and Child Background Characteristics

Because the Program was not allocated to project schools (and not to comparison schools) randomly, it is important to assess whether the two groups are comparable. We did this by examining school and child background data for both states (Chhattisgarh and Uttarakhand) together and for each state separately. (See Tables D.1 and D.2 in *Appendix D,* Tables E.1 and E.2 in *Appendix E* and Tables F.1 and F.2 in *Appendix F* for more details.)

For the two states together, analysis of all observable school level background data (e.g. school location, total enrollment in Grade 2, attendance on the day of the assessment in Grade 2, number of primary-grade teachers, pupil-teacher ratio, mono- versus multi-grade classroom, and time in minutes allocated for reading instruction class per day) did not reveal any statistically significant difference between project and comparison groups. For children, data were collected on age, gender, whether child speaks Hindi at home, whether child attended pre-school (Anganwadi Centre), whether child gets support in study at home, whether child has television and collection of books for his/her reading at home, and whether child's family takes newspaper at home respectively. It was found that the share of children speaking Hindi at home for the two states together was significantly lower in project schools ($p < 0.05$). For rest of the child background characteristics, no statistically significant difference between project and comparison groups were found for the two states together.

In Chhattisgarh, analysis of all observable school level background data did not reveal any statistically significant difference between project and comparison groups. On child background variables, share of children who attended pre-school (Anganwadi Centre) was significantly higher in project schools in contrast to comparison schools ($p < 0.05$). For rest of the child background characteristics, no statistically significant difference between children from project and comparison groups was found in Chhattisgarh.

In Uttarakhand, no statistically significant difference between project and comparison group schools was found on the observable school background characteristics. On child background variables, share of children who speaks Hindi at home was significantly lower in project schools as compared to comparison schools ($p < 0.05$). For rest of the child background characteristics, no statistically significant difference between from project and comparison groups were found in Uttarakhand.

Statistical comparisons of reading assessment results between project and comparison groups overall and by states took into account these differences in child background characteristics.

3.2 Endline Evaluation Results

3.2.1 Impact of the Program on Reading Skills

Table 3.1 provides an overview of the baseline (August 2016) and endline (Feb-March 2018) assessment results by project and comparison group across each assessment task (see *Appendices E* and *F* for state-specific results). **Overall, children from project schools performed better and experienced significantly greater two-year gains than children from comparison schools.** By the end of Grade 2, children from project schools could correctly read an average of 37 words per minute, while children from comparison could read only 18 words correctly per minute. On reading comprehension, children from project schools could correctly answer an average of 2.3 questions correctly (out of 5), compared to just 1.3 questions answered correctly by comparison school children. **Additionally, children in project schools made multifold and statistically significant improvement from baseline**

to endline on all reading assessment tasks when compared with children in comparison schools. Two-year gains on reading skills made by children in project group was 1.4- to 2-times greater than their comparison school counterparts, with maximum gains observed in case of oral reading fluency ($p < 0.001$).

TABLE 3.1: Reading Assessment Results at Baseline and Endline

Assessment Task	Group	Baseline (*with start-of-Grade 1 children*)			Endline (*with end-of-Grade 2 Children*)			2-Year Gains	Difference between 2-Year Gains in Project and Comparison Groups[†]
		n	Mean	SD	n	Mean	SD		
Letter sounding fluency (letters per minute)	Project	607	11.1	14.0	481	54.3	21.3	+43.2	12.6***
	Comparison	509	7.7	11.7	459	38.7	57.5	+31.0	
Non-word reading (non-words per minute)	Project	607	1.2	3.7	481	17.8	10.1	+16.6	6.3***
	Comparison	509	0.6	2.3	459	10.9	14.7	+10.3	
Oral reading fluency (correct words per minute)	Project	607	0.9	3.0	481	36.7	25.0	+35.8	18.4***
	Comparison	509	0.5	2.2	457	17.9	21.9	+17.5	
Reading comprehension (number of questions answered correctly)	Project	607	0.1	0.5	481	2.3	1.4	+2.2	1.0***
	Comparison	509	0.1	0.4	457	1.3	1.5	+1.2	

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

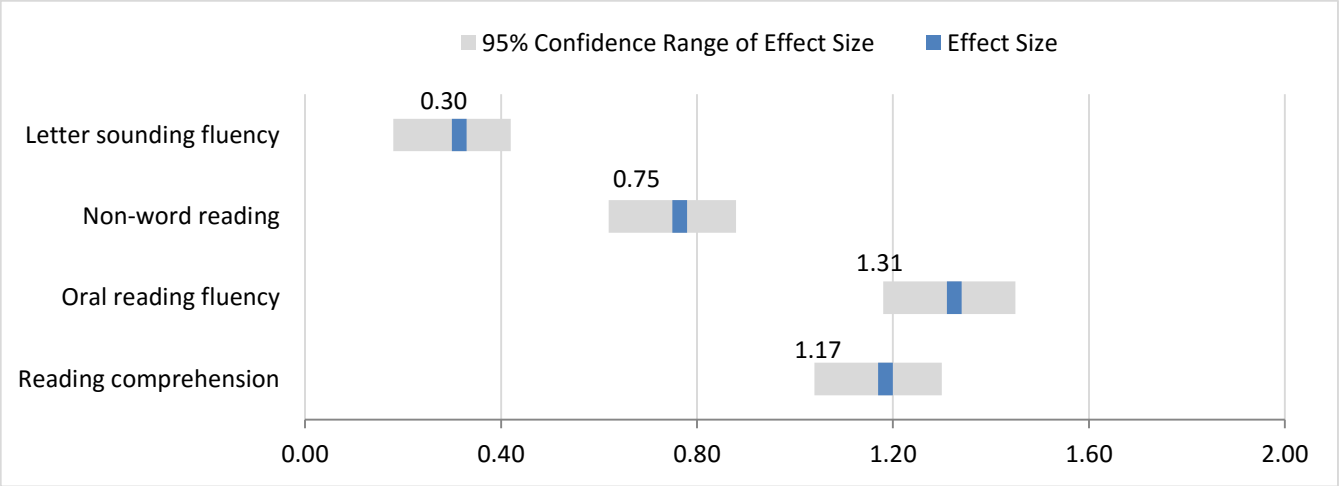
[†] *Difference between 2-Year Gains in Project and Comparison Groups:* This reports the coefficient of the variable on interaction between group (project/ comparison) and wave of data collection (baseline/endline) from a regression analysis with score on each reading assessment task as the dependent variable. This regression analysis was conducted with random effects at the school level and includes age, gender, whether child speaks Hindi at home, group, and wave of data collection as covariates. Additionally, versions of passage used during the tests were included in the regression analysis for reporting results for oral reading fluency and reading comprehension, respectively.

State-wise analysis of data revealed that the Literacy Program under partnership approach was effective in both states. However, two-year gains on reading skills made by project school children from Chhattisgarh was noteworthy. Progress made by children in project schools from baseline to endline in Chhattisgarh was 1.7- to 2.4-times greater than the progress made by children in comparison school across the four reading assessment tasks, with maximum gains registered on oral reading fluency test ($p < 0.001$). In Uttarakhand, two-year gains made by project group children was 1.3- to 1.7-times greater than their counterparts in the comparison school on non-word reading, oral reading fluency and reading comprehension test respectively. However, gains made on letter sounding test by project school children from baseline to endline in Uttarakhand was not significantly different from the same made by comparison school children.

Statistically, differences in the gains made by children from baseline to endline between project and comparison schools across different reading assessment tasks can be better examined by estimating the effect sizes (see Figure 3.1 below). 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 difference in gains from baseline to endline between project school children and comparison school children through linear regression analysis that took into account the clustering effect (see *Appendix B: Data Analysis*) and then dividing this difference by the adjusted pooled standard deviation of children's scores. **The effect sizes for the Literacy Program under partnership approach program were large for two of the four reading assessment tasks, with the largest effect size of 1.31 estimated for oral reading fluency.** Of note, the effect sizes for the Program were higher in Chhattisgarh than Uttarakhand on most of the tasks. In fact, in Chhattisgarh, the effect sizes on all four tasks were

large (more than 0.80). In Uttarakhand, the large effect size of the Program was found in case of oral reading fluency and reading comprehension tests only (see *Appendices E* and *F*).

FIGURE 3.1: Effect Sizes across Reading 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 correctly 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 the Literacy Program under partnership approach in the two states has progressed to achieve the goal by the end of Grade 2. **Endline data from the two states together shows that 32 percent of children in project schools after two years of intervention met or exceeded the fluency benchmark of 45 words read correctly per minute by the end of Grade 2 (compared to only 12 percent children in comparison school).** The difference between project and comparison groups on this benchmark was statistically significant ($p < 0.001$). In Chhattisgarh and Uttarakhand, the shares of end-of-Grade 2 children in project schools who met or exceed the fluency benchmark were 31 percent (compared to just six percent in comparison schools) and 35% (compared to 24% in comparison schools) respectively. However, only in Chhattisgarh, the difference between the groups on this indicator was statistically significant ($p < 0.001$).

Other benchmark indicators that reflect progress include the proportion of children answering at least three and at least four questions correctly on the reading comprehension task. As shown in Figure 3.3, **46 percent project school children could correctly answer at least three comprehension questions correctly (compared to 22 percent of comparison school children) and 23 percent of project school children could correctly answer at least four comprehension questions correctly (compared to just 11 percent of comparison school children) by the end of Grade 2.** These differences were statistically significant (p <0.001). In Chhattisgarh, 44 percent and 20 percent of children in project schools could answer at least three and at least four of the five reading comprehension questions correctly by the end of Grade 2. The corresponding shares in comparison schools were 15% and five percent only; and the differences between the two groups on both of these benchmarks were statistically significant (p <0.001) in Chhattisgarh. In project schools of Uttarakhand, 51 percent (34 percent in comparison school) and 30 percent (20 percent in comparison school) of children could answer at least three and at least four of the five reading comprehension questions correctly by the end of Grade 2. On both of these

benchmarks, the differences between project and comparison groups in Uttarakhand were statistically significant ($p < 0.05$).

FIGURE 3.2: Distribution of Oral Reading Fluency Scores at Endline

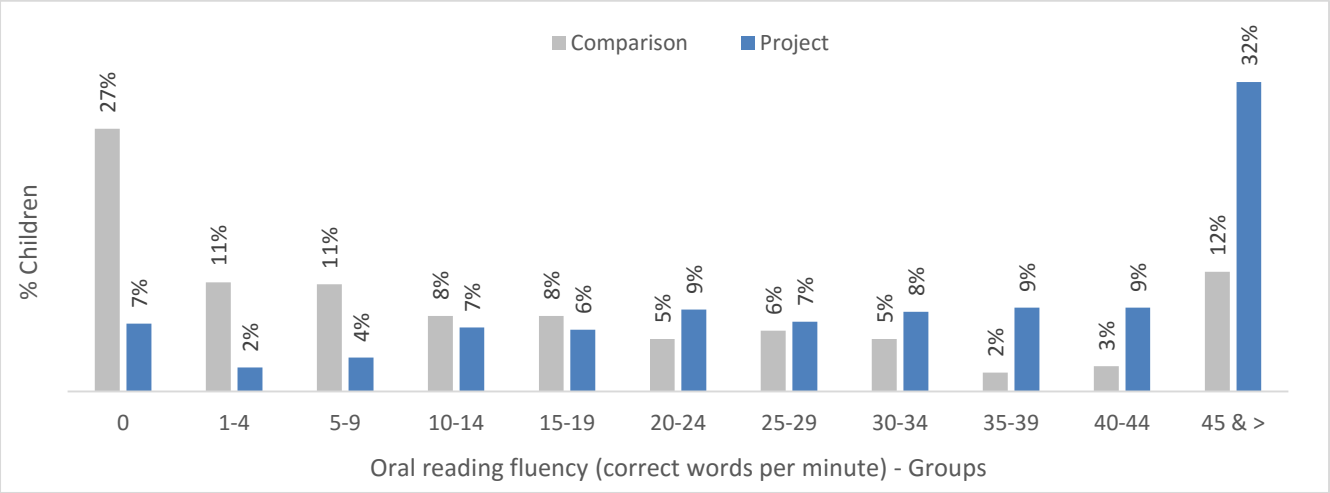
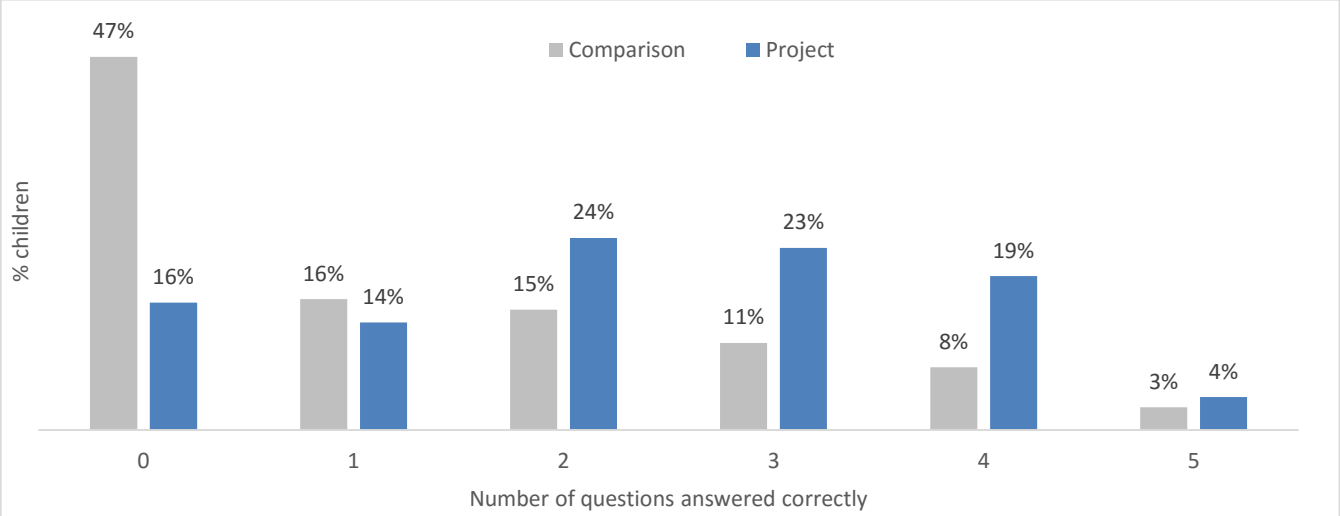


FIGURE 3.3: Distribution of Reading Comprehension Scores at Endline



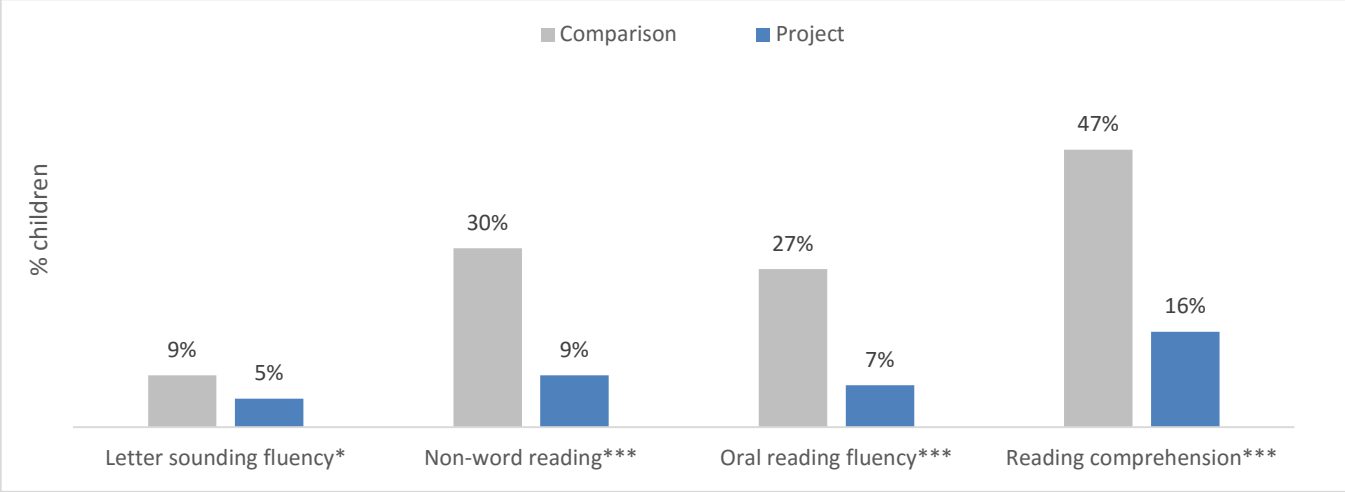
3.2.3 Prevalence of Zero Score

Analysis of zero scores provides another view of how children performed, with a focus on children with the lowest achievement. In the EGRA, 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 of the first pre-decided number of items for a particular assessment task (also known as a discontinued task)[5]. Zero scores on tasks show the subset of children who can be characterized as nonreaders.

[5] For example, on letter sounding fluency test, if a child could not give correct answer on any of the first 10 letters, the test was discontinued and the child was given 0 score on the test. Similarly, on non-word reading test, the threshold for discontinuity was set at first 5 non-words. For oral reading fluency test, we have used four different version of Hindi passages, each having 60 words, with similar level of difficulty in reading. The threshold for discontinuation for each these passages was slightly different from each other, ranging between first 7 to 11 words across passages.

Figure 3.4 below compares the prevalence of zero scores among end-of-Grade 2 children at the endline between the project schools and comparison schools across four reading assessment tasks in the two states together. **The endline data shows that prevalence of zero scores in both project and comparison schools reduced substantially from the levels in baseline, but across all reading assessment tasks the prevalence were significantly lower in the project group than the comparison group after two years of the Program.** In project schools, the highest proportion of zero scores at endline were noted for reading comprehension (16 percent), followed by non-word reading fluency (nine percent), oral reading fluency (seven percent) and letter sounding fluency (five percent). Similar pattern of prevalence of zero scores across four reading assessment tasks were observed for both Chhattisgarh and Uttarakhand at endline. Overall, these data suggest that despite significant improvements some project school children were struggling with higher-order reading skills like blending and reading for understanding.

FIGURE 3.4: Prevalence of Zero Scores at Endline



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

3.3 Comparisons by Gender

Room to Read is also interested in understanding the role gender plays alongside the Literacy Program intervention under the partnership approach. Table 3.2 below provides the assessment results disaggregated by gender for project and comparison school children. Data show that the Program benefitted both boys and girls. Across all reading assessment tasks, both girls and boys from project schools scored higher at endline**. Also, 2-year gains in reading skills made by both boys and girls in project schools were significantly greater than the gains experienced by their respective counterparts in comparison schools** (for boys: $p < 0.05$ on letter sounding fluency and $p < 0.001$ across rest of the three tasks; for girls: $p < 0.001$ across all four tasks)**.** Though both boys and girls benefited from the 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 ($p < 0.001$) and the oral reading fluency ($p < 0.001$) tasks respectively. On letter sounding and reading comprehension tasks, the differences in 2-yeasr gains between project school girls and project school boys were not statistically significant. As shown in Figures 3.5 and 3.6, the effect sizes across all reading tasks were higher for girls than boys. In fact, large effect size (more than 0.80) was for girls were observed for three of the four (non-word reading, oral reading fluency and reading comprehension) assessment tasks. Effect sizes for boys were large (more than 0.80) in case of oral reading fluency and reading comprehension tasks only.

Of note, in Chhattisgarh, differences in 2-year gains across four tasks made by both boys and girls in project schools were significantly greater than the gains experienced by their respective counterparts in comparison schools ($p < 0.001$). In Uttarakhand, differences in 2-year gains made by project school boys were significantly higher than the gains made by comparison school boys on oral reading fluency ($p < 0.001$) and reading comprehension ($p < 0.01$) tasks only. On the other hand, girls in project schools in Uttarakhand made significantly greater gains from baseline to endline on non-word reading ($p < 0.01$), oral reading fluency ($p < 0.001$) and reading comprehension tasks ($p < 0.01$) tasks respectively. Estimated effect size across reading assessment tasks for the two states also reflect the similar picture as presented in *Appendix E* and *Appendix F*.

TABLE 3.2: Reading Assessment Results at Baseline and Endline by Gender

Assessment Task	Group	Baseline (*with start-of-Grade 1 children*)			Endline (*with end-of-Grade 2 children*)			2-Year Gains	Difference between 2-Year Gains in Project and Comparison Groups[†]
		n	Mean	SD	n	Mean	SD		
Letter sounding fluency (letters per minute)	Project Boys	303	9.6	12.6	224	52.6	21.3	+43.0	**D1**: 10.3* **D2**: 14.6*** **D3**: 0.7
	Comparison Boys	239	6.9	11.2	217	40.5	80.5	+33.5	
	Project Girls	304	12.5	15.2	257	55.7	21.2	+43.2	
	Comparison Girls	270	8.4	12.2	242	37.1	22.0	+28.7	
Non-word reading (non-words per minute)	Project Boys	303	0.9	3.2	224	16.2	9.9	+15.2	**D1**: 5.3*** **D2**: 7.3*** **D3**: 2.6***
	Comparison Boys	239	0.5	2.1	217	10.7	18.0	+10.2	
	Project Girls	304	1.5	4.1	257	19.2	10.1	17.7	
	Comparison Girls	270	0.6	2.6	242	11.0	11.0	+10.4	
Oral reading fluency (correct words per minute)	Project Boys	303	0.7	2.6	224	32.3	24.3	+31.6	**D1**: 16.7*** **D2**: 20.6*** **D3**: 8.1***
	Comparison Boys	239	0.4	1.7	215	16.0	20.7	+15.6	
	Project Girls	304	1.1	3.3	257	40.5	25.0	+39.4	
	Comparison Girls	270	0.6	2.5	242	19.7	22.9	+19.1	
Reading comprehension (number of questions answered correctly)	Project Boys	303	0.1	0.4	224	2.2	1.5	+2.1	**D1**: 1.0*** **D2**: 1.0*** **D3**: 0.2
	Comparison Boys	239	0.1	0.3	215	1.2	1.5	+1.1	
	Project Girls	304	0.1	0.5	257	2.4	1.4	+2.2	
	Comparison Girls	270	0.1	0.4	242	1.3	1.5	+1.2	

D1: Difference between 2-Year Gains from baseline to midline between boys in project schools and boys in comparison schools.
D2: Difference between 2-Year Gains from baseline to midline between girls in project schools and girls in comparison schools.
D3: Difference between 2-Year Gains from baseline to midline between girls in project schools and boys in project schools.
Legend of statistical significance of differences between project and comparison schools: ** p < 0.05, ** p < 0.01 and *** p < 0.001*

[†] *Difference between 2-Year Gains in Project and Comparison Groups*: This reports the coefficient of the variable on interaction between type of school (project/ comparison) and round of data collection (baseline/endline) from a regression analysis with score on each reading assessment as the dependent variable for boys and girls separately. This regression analysis was conducted with random effects at the school level and age, whether speaks Hindi at home, type of school and round of data collection as covariates. Additionally, versions of passage used during the tests were included in the regression analysis for reporting results for oral reading fluency and reading comprehension, respectively.

FIGURE 3.5: Effect Sizes across Reading Assessment Tasks for Boys

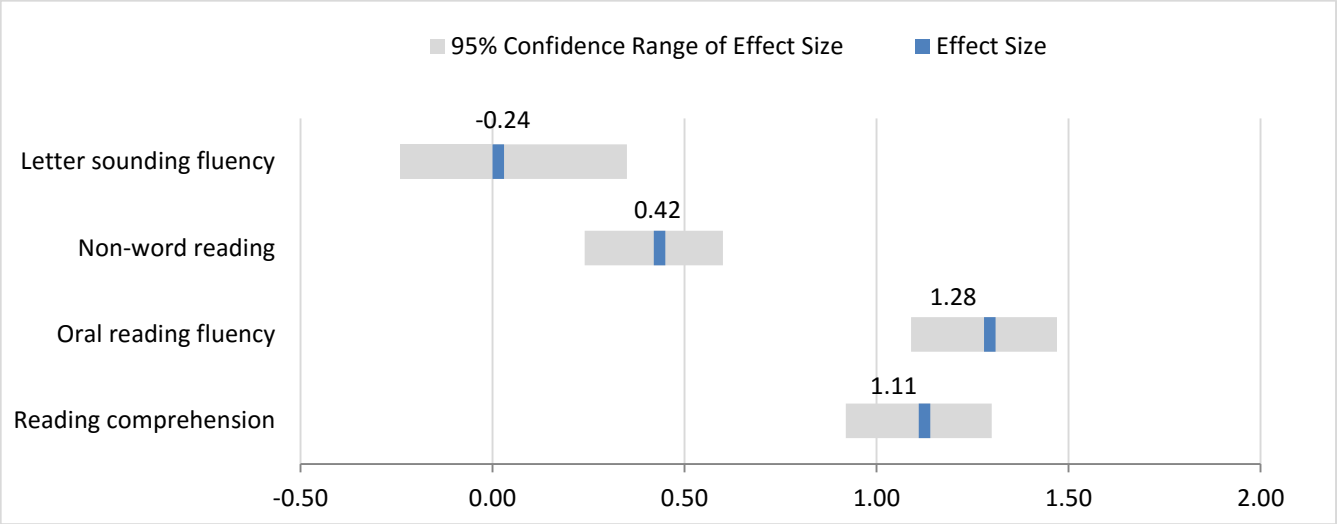
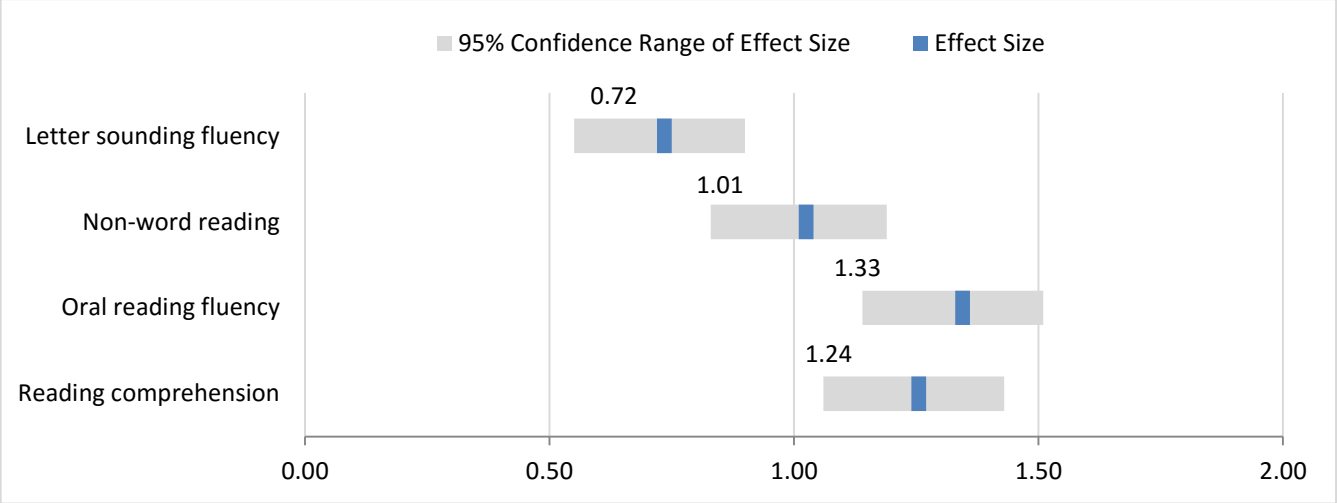


FIGURE 3.6: Effect Sizes across Reading Assessment Tasks for Girls



4 Limitations

Interpretation of the above results should include the limitations of the impact evaluation design. One such limitation of this evaluation was comparability. The validity of impact evaluation results rests on the strength of the assumption that the comparison group, on average, is similar to the project group across all observable and unobservable characteristics that may affect the outcome (reading skills) being evaluated. As explained in section *2.4 Sampling of Schools and Children for Evaluation*, 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 did, in fact, exist. As discussed in section *3.1 School and Child Background Characteristics*, the project and comparison groups did not differ on any of the school background characteristics. Similarly, on majority of the child-background variables, we did not find any statistically significant difference between the children assessed from the project and comparison schools. However, it was found that the share of children speaking Hindi at home in the sample for the two states together and separately in Uttarakhand was significantly lower in project schools ($p < 0.05$). Also, in Chhattisgarh, share of children who attended pre-school (Anganwadi Centre) was significantly higher in project schools in contrast to comparison schools ($p < 0.05$). These characteristics were controlled for while estimating the differences between 2-year gains on reading skills from baseline to endline made by children from project and comparison groups. However, because project group schools were not randomly selected for the Literacy Program under partnership approach in the two states, it may be possible that they differed from comparison schools in ways that the evaluation did not assess or could not observe.

Lastly, the child sample size for reading assessments at both baseline and endline was little lower than what was planned originally. We aimed to collect data from 10 children in each school. However, the achieved sample of children was lower than anticipated number due to high student absenteeism in both project and comparison schools. The sample size for schools (in both project and comparison group), the primary unit of measurement in this impact evaluation, was not affected by this challenge.

5 Conclusion

Findings from the 2018 endline data indicate that the Literacy Program under partnership approach supported by SERI 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 at endline. Most importantly, children in project schools also experienced significantly greater gains in reading skills than children in comparison schools from the beginning of Grade 1 to the end of Grade 2. Of note, the program was similarly effective for boys and girls. Also, the Program was quite effective in both states (Chhattisgarh and Uttarakhand), however, the benefits experienced by project school children in Chhattisgarh was notable.

Despite these results, room for improvement exists. For example, though the relative impact of the program was large, the overall reading fluency levels of project school children (37 correct words per minute, on average) was below our goal of at least 45 correct words per minute by the end of Grade 2. Moreover, seven percent of project school children (and 27 percent of comparison school children) were effectively non-readers, while 16 percent of project school children (and 47 percent of comparison school children) were unable to answer even one reading comprehension question correctly. These results suggest that more can be done to improve the quality and effectiveness of program delivery.

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.

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

Appendix A: Instruction Component of the Hindi Literacy Program

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/ Facilitators employed by Room to Read provide classroom support to teachers throughout the intervention.

During the research and development stage of the instruction component of the Program, Room to Read India country team analyzed 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 oral and silent reading which 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 student workbooks provide daily opportunities to practice the writing skills taught.

Appendix B: Data Analysis[6]

The first aim of the data analysis was to determine if there were statistically significant differences in school and child background variables between the project and comparison groups. For the school background variables, we examined differences between project and comparison group in total enrollment in Grade 2, attendance rate in Grade 2 on the day of the assessment, total number of teachers in primary grade and pupil-teacher ratio by conducting *t* tests and differences in school location (rural vs. urban), type of classroom (mono-grade vs. multi-grade) and duration of reading instruction class per day by categories by conducting chi-square tests. For the child background variables, we examined differences between children from project and comparison schools in age, gender, whether child speaks Hindi at home, whether child attended pre-school (Anganwadi Centre), whether child gets support in study at home, whether child has television and collection of books for his/her reading at home, and whether child's family takes newspaper at home by conducting regression analysis (linear regression for the continuous variables and logistic regression for the categorical variables) with random effects at the school level. The equations included the child background variables as independent variables and group type (project or comparison) as the predictor. The results of these analyses conducted for the two states together appear in *Appendix D* and state-specific analyses appear in *Appendix E* and *Appendix F*.

The primary aim of the data analysis was to determine if children in the project group made greater gains from baseline to endline after two years of intervention than children in the comparison group without any intervention. 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. While analyzing the data together from the two states and separately from Uttarakhand, we have also included age, gender and whether child speaks Hindi at home as the covariates in each regression model. On the other hand, each regression model for Chhattisgarh included age, gender and whether child attended pre-school (Anganwadi Centre) as the covariates.[7] Additionally, versions of Hindi passage used during the tests were included in the regression analysis for reporting results for oral reading fluency and reading comprehension tests in the corresponding regression models for the overall and state-specific analyses.

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 other covariates that are mentioned in the previous paragraph for the overall or states-specific analyses.

---

[6] All data analyses were conducted using Stata statistical software (Stata Corp, 2013).

[7] Age and sex were included in the regression models because of their known effects on children' reading performance. Other covariates like whether child speaks Hindi at home or whether child attended pre-school was included in overall or state-specific analysis because of a significant difference between project and comparison school children on this variable (see *Appendix D, E and F*).

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., the distribution 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 regression equations included the presence or absence of a zero score as the dependent variable, type of school as the predictor, and age and whether child speaks Hindi at home as covariates while analyzing data from the two states together and for Uttarakhand separately. Logistic regression models for Chhattisgarh included the same dependent variable and predictor and age and whether child attended pre-school as covariates. Lastly, logistic regression models for prevalence of zero score in oral reading fluency and reading comprehension tests included version of passage for the test as additional covariates.

We also estimated effect size to demonstrate and compare the magnitude of the impact of the Literacy Program under partnership approach for different outcome measures. Effect size is the 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.

Appendix C: Survey Management

Assessor Training

Sigma Research and Consulting Private Limited, the survey agency hired by Room to Read for 2018 endline EGRA data collection, organized a five-day training workshop for assessors in Dehradun during Feb 1-5, 2018. Members of the Research, Monitoring, and Evaluation (RM&E) unit of Room to Read, India provided the training to the assessors in this workshop with technical support from Room to Read's Instruction Design and Technical Support (IDTS) unit and coordination support from Sigma. It needs to be mentioned here that Room to Read used Computer Aided Personal Interview (CAPI) method for collection of EGRA data through tablet with the help Tangerine software[8].The five-day training workshop included three days of desk-based training where the focus was on explanation of the four EGRA sub-tests, administration processes, use of Tangerine for EGRA data collection through tablet, demonstration by experts, and mock tests. The training also included a two-day exposure visit to a non-sample school for practice with children and a half day of reflection. At the end of the workshop, an inter-rater reliability test was conducted to select the best 60 assessors (out of all participants in the workshop) on the basis of their test score, and 30 of them were assigned the role of supervisors.

Data Collection

2018 Endline data collection was conducted almost simultaneously in the two states (Chhattisgarh and Uttarakhand). Data collection was first initiated in the Uttarakhand on February 6, 2018 and the survey in 70 schools (36 project and 34 comparison) was completed by February 23, 2018. In Chhattisgarh, data collection in 76 school (38 project and 38 comparison) took place during February 18-26, 2018. In Uttarakhand, specifically in Champawat (the district where the project schools were located), duration of data collection was longer because of the hilly terrain and hard to reach schools which had slowed the pace of the survey. In each state, data collection was conducted by multiple teams, and each team was comprised of one assessor and one supervisor. The role of assessor was to assess the children whereas the supervisors were assigned with the tasks of field management, ensuring the compliance with assessment administration process, and maintaining data quality at field by jointly assessing the first three children during data collection.

[8] Tangerine, a software for EGRA data collection, was developed by RTI International. Room to Read has authorised access to use Tangerine.

Appendix D: School and Child Background Characteristics

TABLE D.1: Background Characteristics of Sample School at Endline*

	Project		Comparison	
	n	Mean (SD) or %	n	Mean (SD) or %
School	74	-	72	-
State				
Chhattisgarh	38	51%	38	53%
Uttarakhand	36	49%	34	47%
Location				
Rural	70	95%	66	92%
Urban	4	5%	6	8%
Enrollment in Grade 2	74	14 (11)	72	14 (14)
Attendance in Grade 2 on day of endline assessment (%)	74	69 (22)	72	75 (22)
Total teachers in primary grades	74	3 (2)	72	3 (2)
Pupil-teacher ratio	74	5 (2)	72	4 (3)
Classroom type				
Mono-grade	40	54%	29	40%
Multi-grade	34	46%	43	60%
Designated minutes of reading instruction per day				
less 40 min	7	9%	13	18%
40-50 min	30	41%	27	38%
50-60 min	27	36%	20	28%
more than 60 min	10	14%	12	17%

*Differences between project and comparison schools on the background characteristics were not statistically significant

TABLE D.2: Background Characteristics of Sample Children

	Project		Comparison	
	n	Mean (SD) or %	n	Mean (SD) or %
Children - Baseline (2016)	607	-	509	-
Children - Endline (2018)	481	-	460	-
State - Baseline (2016)				
Chhattisgarh	386	64%	369	72%
Uttarakhand	221	36%	140	28%
State - Endline (2018)				
Chhattisgarh	302	63%	289	63%
Uttarakhand	179	37%	171	37%
Gender - Baseline (2016)				
Boys	303	50%	239	47%
Girls	304	50%	270	53%
Gender - Endline (2018)				
Boys	224	47%	218	47%
Girls	257	53%	242	53%
Age in years - Endline (2018)	481	7.5 (0.8)	460	7.4 (0.9)
Speaks Hindi at home - Endline (2018)*	188	39%	238	52%
Attended pre-school (Anganwadi Centre) - Endline (2018)	364	76%	389	85%
Has TV at home - Endline (2018)	352	73%	353	77%
Has collection of books at home - Endline (2018)	267	56%	279	61%
Takes newspaper at home - Endline (2018)	359	75%	374	81%
Gets support in study at home - Endline (2018)	98	20%	125	27%

Legend of statistical significance of the difference between project and comparison schools on child background characteristics: ** p < 0.05, ** p < 0.01 and *** p < 0.001*

TABLE D.3: Correlations between the Reading Assessment Scores at Endline

	Letter sounding fluency (letters per minute)	Non-word reading (non-words per minute)	Oral reading fluency (correct words per minute)	Reading comprehension (number of questions answered correctly)
Letter sounding fluency (letters per minute)	1.000			
Non-word reading (non-words per minute)	0.785***	1.000		
Oral reading fluency (correct words per minute)	0.735***	0.881***	1.000	
Reading comprehension (number of questions answered correctly)	0.617***	0.688***	0.694***	1.000

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

Appendix E: Results Specific to Chhattisgarh

TABLE E.1: Background Characteristics of Sample School at Endline in Chhattisgarh*

	Project		Comparison	
	n	Mean (SD) or %	n	Mean (SD) or %
School	38	-	38	-
Location				
Rural	34	89%	35	92%
Urban	4	11%	3	8%
Enrollment in Grade 2	38	19 (12)	38	20 (17)
Attendance in Grade 2 on day of endline assessment (%)	38	64 (19)	38	71 (18)
Total teachers in primary grades	38	4 (2)	38	4 (2)
Pupil-teacher ratio	38	5 (2)	38	5 (2)
Classroom type				
Mono-grade	29	76%	24	63%
Multi-grade	9	24%	14	37%
Designated minutes of reading instruction per day				
less 40 min	0	0%	2	5%
40-50 min	15	39%	18	47%
50-60 min	17	45%	11	29%
more than 60 min	6	16%	7	18%

* Differences between project and comparison schools on background characteristics were nor statistically significant

TABLE E.2: Background Characteristics of Sample Children in Chhattisgarh

	Project		Comparison	
	n	Mean (SD) or %	n	Mean (SD) or %
Children - Baseline (2016)	386	-	369	-
Children - Endline (2018)	302	-	289	-
Gender - Baseline (2016)				
Boys	198	51%	180	49%
Girls	188	49%	189	51%
Gender - Endline (2018)				
Boys	141	47%	140	48%
Girls	161	53%	149	52%
Age in years - Endline (2018)	302	7.4 (0.7)	289	7.4 (0.9)
Speaks Hindi at home - Endline (2018)	108	36%	125	43%
Attended pre-school (Anganwadi Centre) - Endline (2018)*	257	85%	224	78%
Has TV at home - Endline (2018)	253	84%	245	85%
Has collection of books at home - Endline (2018)	154	51%	169	58%
Takes newspaper at home - Endline (2018)	46	15%	63	22%
Gets support in study at home - Endline (2018)	224	74%	221	76%

Legend of statistical significance of the difference between project and comparison schools on child background characteristics: ** p < 0.05, ** p < 0.01 and *** p < 0.001*

TABLE E.3: Correlations between the Reading Assessment Scores at Endline in Chhattisgarh

	Letter sounding fluency (letters per minute)	Non-word reading (non-words per minute)	Oral reading fluency (correct words per minute)	Reading comprehension (number of questions answered correctly)
Letter sounding fluency (letters per minute)	1.000			
Non-word reading (non-words per minute)	0.721***	1.000		
Oral reading fluency (correct words per minute)	0.728***	0.886***	1.000	
Reading comprehension (number of questions answered correctly)	0.581***	0.663***	0.678***	1.000

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

TABLE E.4: Reading Assessment Results at Baseline and Endline in Chhattisgarh

Assessment Task	Group	Baseline (*with start-of-Grade 1 children*)			Endline (*with end-of-Grade 2 Children*)			2-Year Gains	Difference between 2-Year Gains in Project and Comparison Groups†
		n	Mean	SD	n	Mean	SD		
Letter sounding fluency (letters per minute)	Project	386	6.2	9.5	302	53.1	20.6	46.9	19.3***
	Comparison	369	6.0	10.0	289	33.5	19.6	27.6	
Non-word reading (non-words per minute)	Project	386	0.2	1.1	302	16.9	9.7	16.7	7.9***
	Comparison	369	0.2	1.1	289	8.8	8.8	8.7	
Oral reading fluency (correct words per minute)	Project	386	0.2	1.1	302	34.8	24.2	34.6	20.0***
	Comparison	369	0.1	1.1	289	14.4	16.5	14.2	
Reading comprehension (number of questions answered correctly)	Project	386	0.0	0.2	302	2.2	1.4	2.2	1.2***
	Comparison	369	0.0	0.3	289	1.0	1.2	1.0	

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

† *Difference between 2-Year Gains in Project and Comparison Groups:* This reports the coefficient of the variable on interaction between group (project/ comparison) and wave of data collection (baseline/endline) from a regression analysis with score on each reading assessment task as the dependent variable. This regression analysis was conducted with random effects at the school level and includes age, gender, whether child attended pre-school, group, and wave of data collection as covariates. Additionally, versions of passage used during the tests were included in the regression analysis for reporting results for oral reading fluency and reading comprehension, respectively.

FIGURE E.1: Effect Sizes across Reading Assessment Tasks in Chhattisgarh

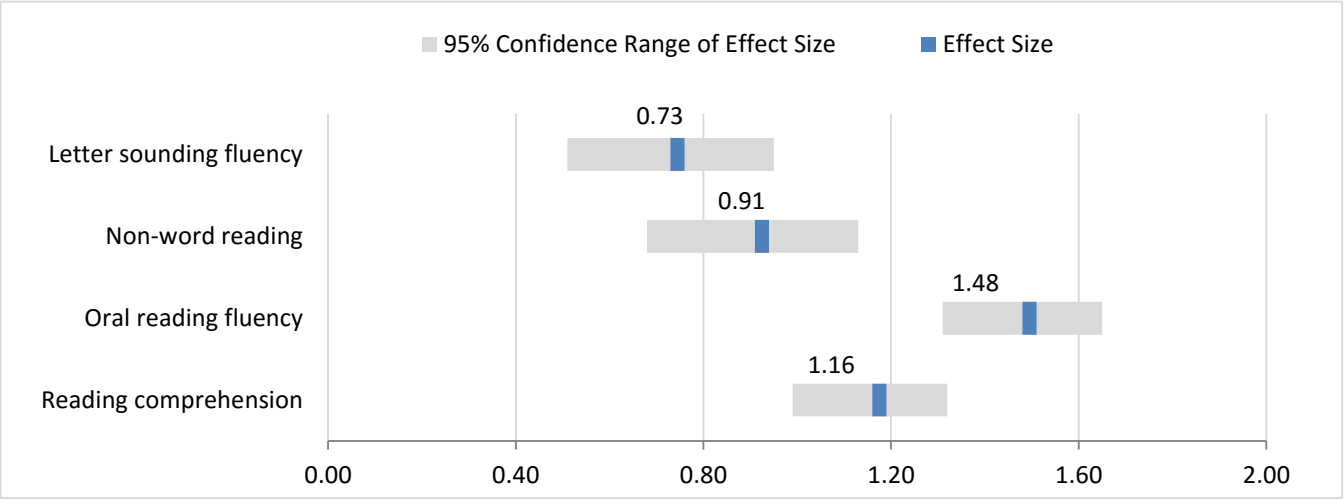


FIGURE E.2: Distribution of Oral Reading Fluency Scores at Endline in Chhattisgarh

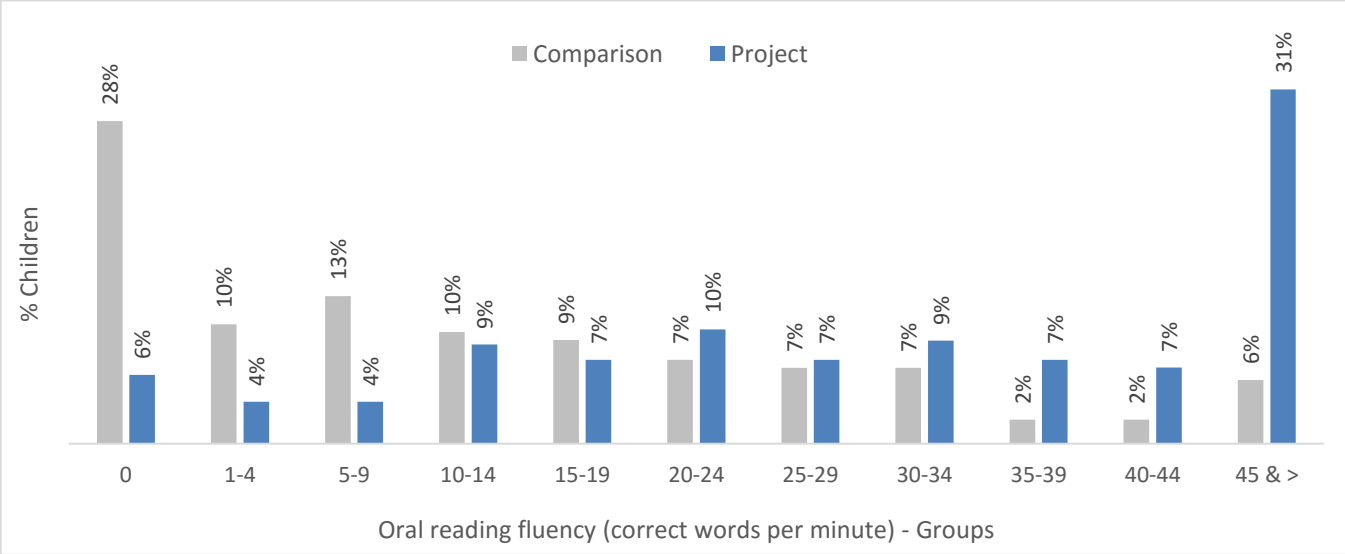


FIGURE E.3: Distribution of Reading Comprehension Scores at Endline in Chhattisgarh

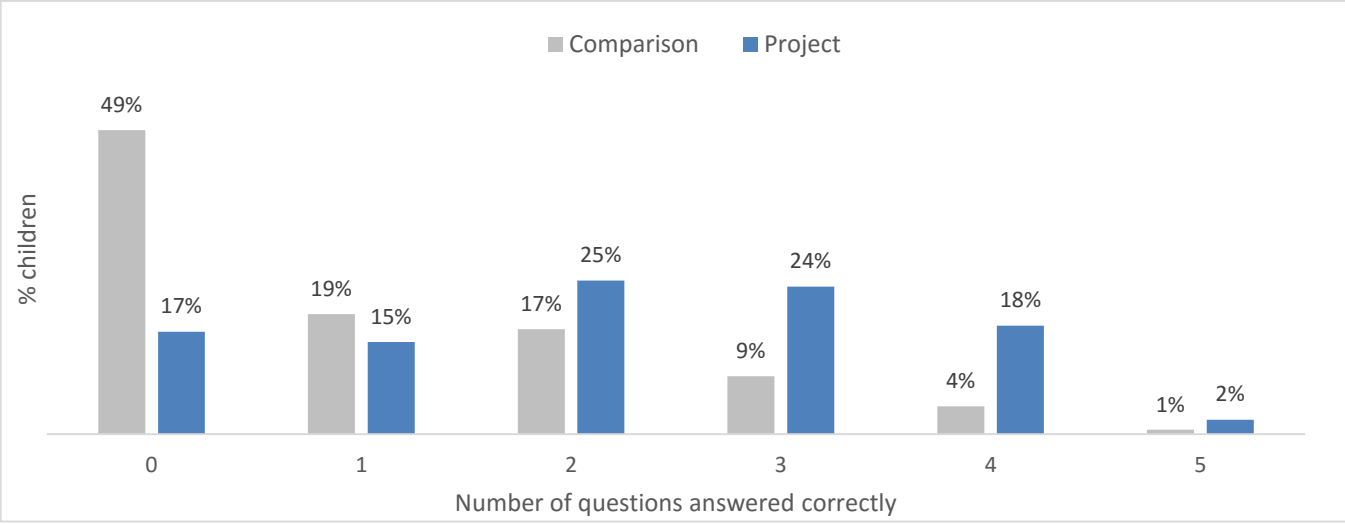
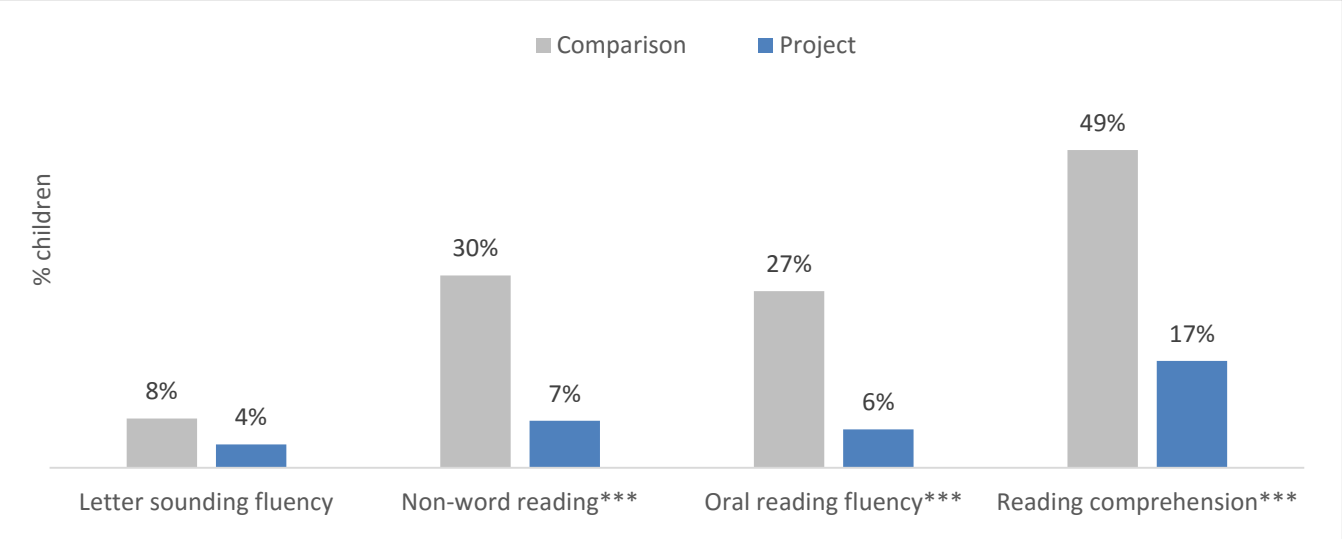


FIGURE E.4: Prevalence of Zero Scores at Endline in Chhattisgarh



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

TABLE E.5: Reading Assessment Results at Baseline and Endline by Gender in Chhattisgarh

Assessment Task	Group	Baseline (*with start-of-Grade 1 children*)			Endline (*with end-of-Grade 2 children*)			2-Year Gains	Difference between 2-Year Gains in Project and Comparison Groups[†]
		n	Mean	SD	n	Mean	SD		
Letter sounding fluency (letters per minute)	Project Boys	198	5.8	9.4	141	51.1	21.6	+45.3	**D1**: 18.5*** **D2**: 20.0*** **D3**: 3.2
	Comparison Boys	180	5.3	9.3	140	32.0	19.8	+26.8	
	Project Girls	188	6.6	9.6	161	55.0	19.6	+48.4	
	Comparison Girls	189	6.7	10.5	149	34.9	19.3	+28.3	
Non-word reading (non-words per minute)	Project Boys	198	0.2	1.3	141	15.1	9.7	+14.9	**D1**: 7.0*** **D2**: 8.7*** **D3**: 3.4***
	Comparison Boys	180	0.2	1.3	140	8.1	8.3	+7.9	
	Project Girls	188	0.1	0.9	161	18.4	9.5	+18.3	
	Comparison Girls	189	0.2	0.9	149	9.5	9.2	+9.4	
Oral reading fluency (correct words per minute)	Project Boys	198	0.2	1.1	141	30.4	23.6	+30.2	**D1**: 17.4*** **D2**: 22.7*** **D3**: 8.5***
	Comparison Boys	180	0.1	0.8	140	12.9	15.6	+12.8	
	Project Girls	188	0.1	1.1	161	38.6	24.2	+38.5	
	Comparison Girls	189	0.2	1.3	149	15.8	17.2	+15.6	
Reading comprehension (number of questions answered correctly)	Project Boys	198	0.0	0.3	141	2.1	1.4	+2.0	**D1**: 1.1*** **D2**: 1.2*** **D3**: 0.2
	Comparison Boys	180	0.0	0.3	140	1.0	1.2	+0.9	
	Project Girls	188	0.0	0.1	161	2.3	1.4	+2.3	
	Comparison Girls	189	0.0	0.2	149	1.1	1.2	+1.0	

D1: Difference between 2-Year Gains from baseline to midline between boys in project schools and boys in comparison schools.
D2: Difference between 2-Year Gains from baseline to midline between girls in project schools and girls in comparison schools.
D3: Difference between 2-Year Gains from baseline to midline between girls in project schools and boys in project schools.
Legend of statistical significance of differences between project and comparison schools: ** p < 0.05, ** p < 0.01 and *** p < 0.001*
[†] *Difference between 2-Year Gains in Project and Comparison Groups*: This reports the coefficient of the variable on interaction between type of school (project/ comparison) and round of data collection (baseline/endline) from a regression analysis with score on each reading assessment as the dependent variable for boys and girls separately. This regression analysis was conducted with random effects at the school level and age, whether child attended pre-school, type of school and round of data collection as covariates. Additionally, versions of passage used during the tests were included in the regression analysis for reporting results for oral reading fluency and reading comprehension, respectively.

FIGURE E.5: Effect Sizes across Reading Assessment Tasks for Boys in Chhattisgarh

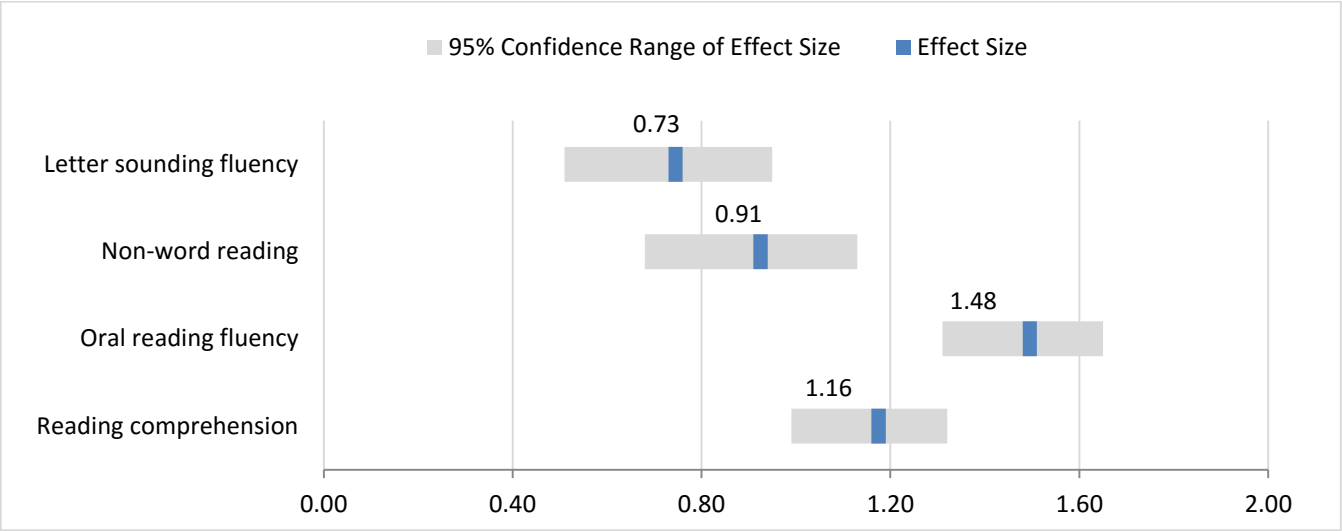
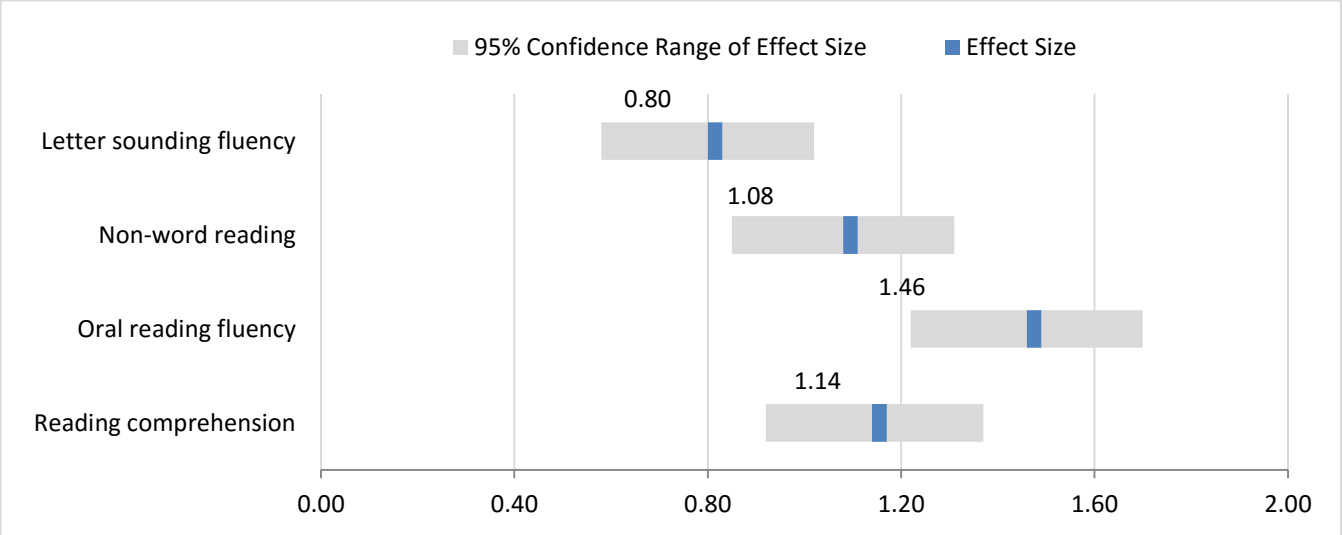


FIGURE E.6: Effect Sizes across Reading Assessment Tasks for Girls in Chhattisgarh



Appendix F: Results Specific to Uttarakhand

TABLE F.1: Background Characteristics of Sample School at Endline in Uttarakhand*

	Project		Comparison	
	n	Mean (SD) or %	n	Mean (SD) or %
School	**36**		**34**	
District				
Champawat	36	-	0	-
Nainital	0	-	34	-
Location				
Rural	36	100%	31	91%
Urban	0	0%	3	9%
Enrollment in Grade 2	36	9 (6)	34	8 (6)
Attendance in Grade 2 on day of endline assessment (%)	36	74 (24)	34	80 (26)
Total teachers in primary grades	36	2 (1)	34	2 (1)
Pupil-teacher ratio	36	4 (2)	34	4 (3)
Classroom type				
Mono-grade	11	31%	5	15%
Multi-grade	25	69%	29	85%
Designated minutes of reading instruction per day				
less 40 min	7	19%	11	32%
40-50 min	15	42%	9	26%
50-60 min	10	28%	9	26%
more than 60 min	4	11%	5	15%

* Differences between project and comparison schools on background characteristics were nor statistically significant.

TABLE F.2: Background Characteristics of Sample Children in Uttarakhand

	Project		Comparison	
	n	Mean (SD) or %	n	Mean (SD) or %
Children - Baseline (2016)	221		140	
Children - Endline (2018)	179		171	
Gender - Baseline (2016)				
Boys	105	48%	59	42%
Girls	116	52%	81	58%
Gender - Endline (2018)				
Boys	83	46%	78	46%
Girls	96	54%	93	54%
Age in years - Endline (2018)	179	7.5 (1.0)	171	7.4 (1.0)
Speaks Hindi at home - Endline (2018)*	80	45%	113	66%
Attended pre-school (Anganwadi Centre) - Endline (2018)	132	74%	140	82%
Has TV at home - Endline (2018)	99	55%	108	63%
Has collection of books at home - Endline (2018)	113	63%	110	64%
Takes newspaper at home - Endline (2018)	52	29%	62	36%
Gets support in study at home - Endline (2018)	150	84%	138	81%

Legend of statistical significance of the difference between project and comparison schools on child background characteristics: ** p < 0.05, ** p < 0.01 and *** p < 0.001*

TABLE F.3: Correlations between the Reading Assessment Scores at Endline in Uttarakhand

	Letter sounding fluency (letters per minute)	Non-word reading (non-words per minute)	Oral reading fluency (correct words per minute)	Reading comprehension (number of questions answered correctly)
Letter sounding fluency (letters per minute)	1.000			
Non-word reading (non-words per minute)	0.840***	1.000		
Oral reading fluency (correct words per minute)	0.738***	0.870***	1.000	
Reading comprehension (number of questions answered correctly)	0.653***	0.706***	0.700***	1.000

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

TABLE F.4: Reading Assessment Results at Baseline and Endline in Uttarakhand

Assessment Task	Group	Baseline (*with start-of-Grade 1 children*)			Endline (*with end-of-Grade 2 Children*)			2-Year Gains	Difference between 2-Year Gains in Project and Comparison Groups†
		n	Mean	SD	n	Mean	SD		
Letter sounding fluency (letters per minute)	Project	221	19.6	16.4	179	56.1	22.3	36.5	1.9
	Comparison	140	12.2	14.6	170	47.4	90.5	35.2	
Non-word reading (non-words per minute)	Project	221	3.0	5.4	179	19.3	10.6	16.3	3.5*
	Comparison	140	1.6	3.9	170	14.3	20.9	12.8	
Oral reading fluency (correct words per minute)	Project	221	2.1	4.4	179	39.9	25.9	37.8	15.4***
	Comparison	140	1.3	3.6	168	24.1	28.0	22.8	
Reading comprehension (number of questions answered correctly)	Project	221	0.3	0.7	179	2.5	1.5	2.2	0.7***
	Comparison	140	0.2	0.5	168	1.6	1.7	1.5	

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

† *Difference between 2-Year Gains in Project and Comparison Groups:* This reports the coefficient of the variable on interaction between group (project/ comparison) and wave of data collection (baseline/endline) from a regression analysis with score on each reading assessment task as the dependent variable. This regression analysis was conducted with random effects at the school level and includes age, gender, whether child speaks Hindi at home, group, and wave of data collection as covariates. Additionally, versions of passage used during the tests were included in the regression analysis for reporting results for oral reading fluency and reading comprehension, respectively.

FIGURE F.1: Effect Sizes across Reading Assessment Tasks in Uttarakhand

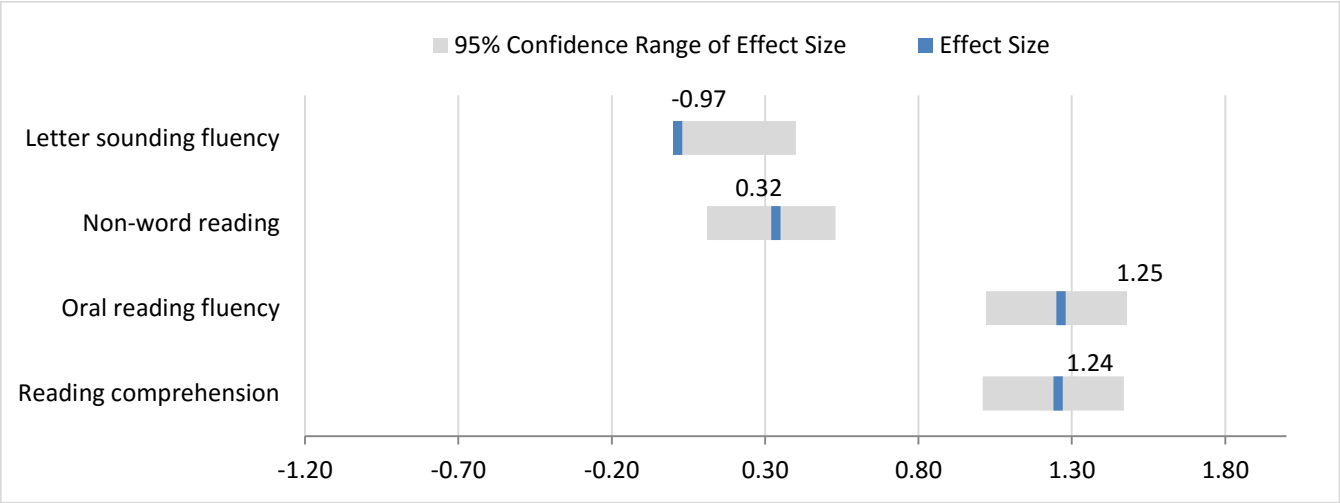


FIGURE F.2: Distribution of Oral Reading Fluency Scores at Endline in Uttarakhand

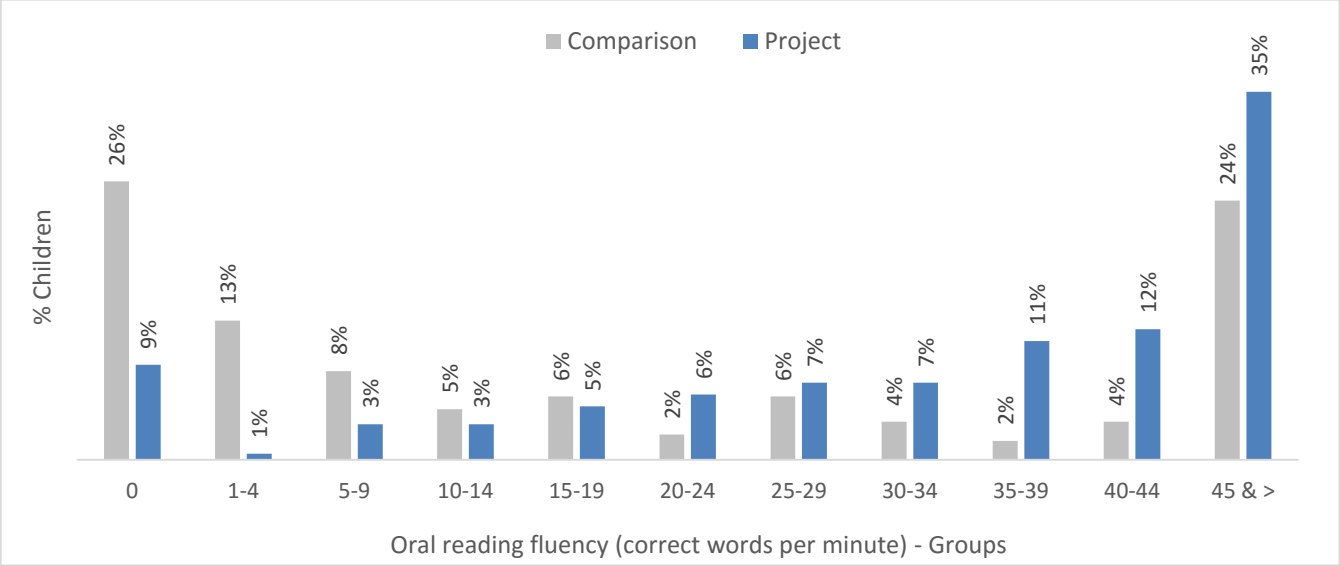


FIGURE F.3: Distribution of Reading Comprehension Scores at Endline in Uttarakhand

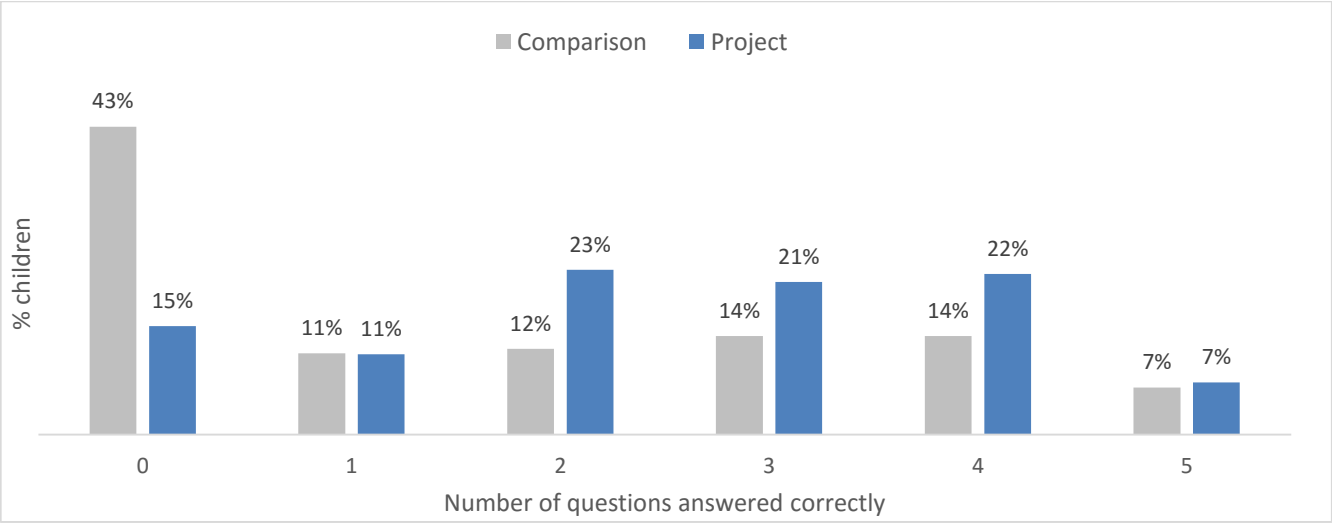


FIGURE F.4: Prevalence of Zero Scores at Endline in Uttarakhand

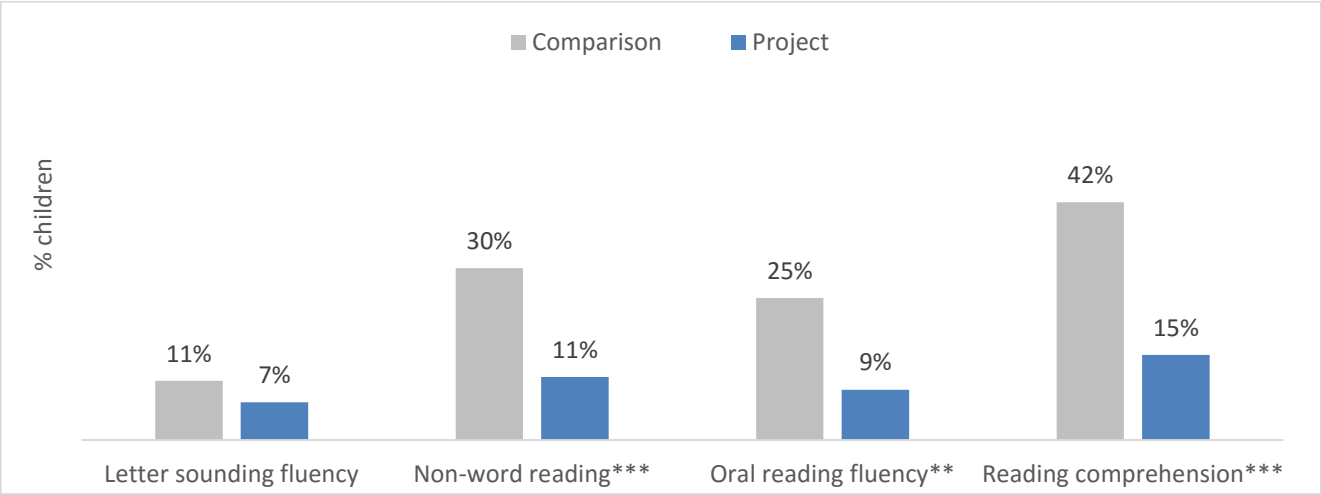


TABLE F.5: Reading Assessment Results at Baseline and Endline by Gender in Uttarakhand

Assessment Task	Group	Baseline (*with start-of-Grade 1 children*)			Endline (*with end-of-Grade 2 children*)			2-Year Gains	Difference between 2-Year Gains in Project and Comparison Groups[†]
		n	Mean	SD	n	Mean	SD		
Letter sounding fluency (letters per minute)	Project Boys	105	16.8	14.6	83	55.3	20.8	+38.5	**D1**: -2.8 **D2**: 6.0 **D3**: -3.2
	Comparison Boys	59	12.0	14.6	77	55.7	131.6	+43.7	
	Project Girls	116	22.1	17.5	96	56.9	23.7	+34.8	
	Comparison Girls	81	12.4	14.7	93	40.6	25.4	+28.2	
Non-word reading (non-words per minute)	Project Boys	105	2.2	4.8	83	17.9	10.0	+15.7	**D1**: 2.0 **D2**: 5.2** **D3**: 1.1
	Comparison Boys	59	1.4	3.5	77	15.4	27.6	+14.0	
	Project Girls	116	3.7	5.9	96	20.5	11.0	+16.7	
	Comparison Girls	81	1.7	4.3	93	13.5	13.0	+11.8	
Oral reading fluency (correct words per minute)	Project Boys	105	1.6	4.0	83	35.6	25.2	+34.0	**D1**: 14.3*** **D2**: 17.6*** **D3**: 7.4*
	Comparison Boys	59	1.1	3.0	75	21.8	27.1	+20.7	
	Project Girls	116	2.5	4.8	96	43.6	26.1	+41.1	
	Comparison Girls	81	1.4	4.0	93	25.9	28.8	+24.5	
Reading comprehension (number of questions answered correctly)	Project Boys	105	0.2	0.6	83	2.4	1.6	+2.2	**D1**: 0.7** **D2**: 0.7** **D3**: 0.0
	Comparison Boys	59	0.2	0.4	75	1.7	1.7	+1.5	
	Project Girls	116	0.3	0.8	96	2.5	1.4	+2.2	
	Comparison Girls	81	0.1	0.6	93	1.6	1.7	+1.5	

D1: Difference between 2-Year Gains from baseline to midline between boys in project schools and boys in comparison schools.
D2: Difference between 2-Year Gains from baseline to midline between girls in project schools and girls in comparison schools.
D3: Difference between 2-Year Gains from baseline to midline between girls in project schools and boys in project schools.
Legend of statistical significance of differences between project and comparison schools: ** p < 0.05, ** p < 0.01 and *** p < 0.001*

[†] *Difference between 2-Year Gains in Project and Comparison Groups*: This reports the coefficient of the variable on interaction between type of school (project/ comparison) and round of data collection (baseline/endline) from a regression analysis with score on each reading assessment as the dependent variable for boys and girls separately. This regression analysis was conducted with random effects at the school level and age, whether child speaks Hindi at home, type of school and round of data collection as covariates. Additionally, versions of passage used during the tests were included in the regression analysis for reporting results for oral reading fluency and reading comprehension, respectively.

FIGURE F.5: Effect Sizes across Reading Assessment Tasks for Boys in Uttarakhand

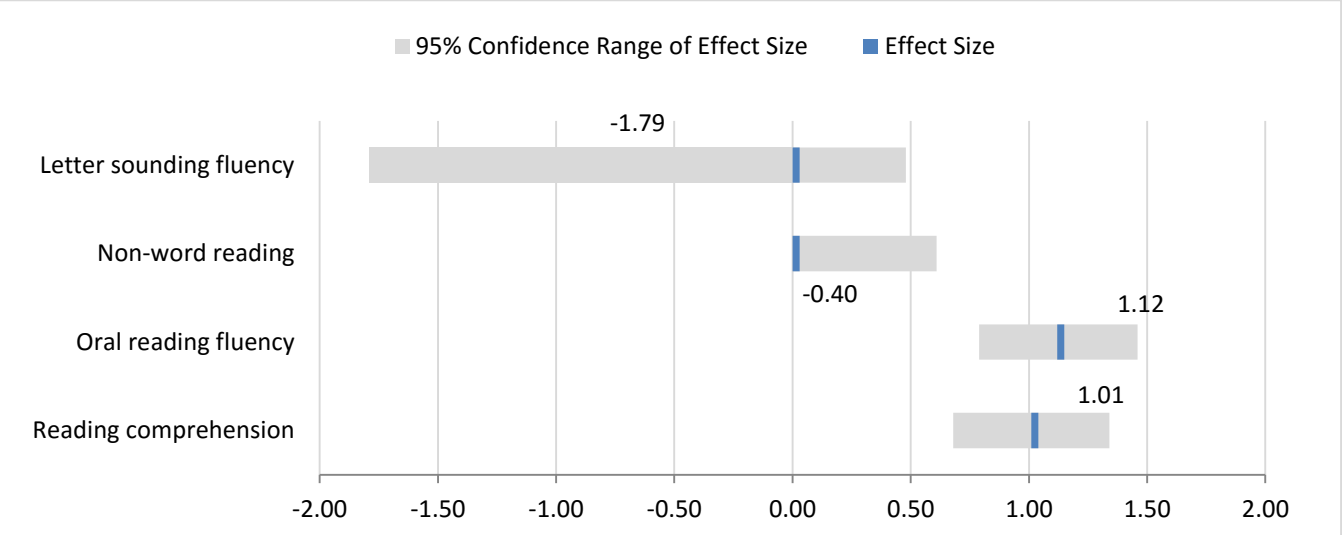


FIGURE F.6: Effect Sizes across Reading Assessment Tasks for Girls in Uttarakhand